ACTIVITY _ ACCESSIBILITY MODELS OF TRIP GENERATION

MAY 1969 - NUMBER 10

BY

T. Z. NAKKASH

JHRP

JOINT HIGHWAY RESEARCH PROJECT

PURDUE UNIVERSITY AND

INDIANA STATE HIGHWAY COMMISSION

Progress Report

ACTIVITY - ACCESSIBILITY MODELS OF TRIP GENERATION

To: J. F. McLaughlin, Director
Joint Highway Research Project

May 6, 1969

Project: C-36-69E

From: H. L. Michael, Associate Director
Joint Highway Research Project

File: 3-7-5

The attached report is submitted as a Progress Report on the HPR-1 (6), Part I project "An Investigation of Major Aspects of the Urban Transportation Planning Process." The report is the first one on Part IV "The Trip Generation Process" and completes the approved research on that Part of the project. The research report is titled "Activity - Accessibility Models of Trip Generation" and has been authored by Mr. Tammam Z. Nakkash. The research has been performed under the direction of Professor W. L. Grecco, Research Engineer on our staff. Mr. Nakkash has also used the research and the report as his thesis for the Ph.D. degree from Purdue University.

The purpose of this research was to examine the effect of activity - accessibility variables on trip generation. The results indicate that relative accessibility of zones of the travel study area should be considered in estimating future trips.

The research has been performed with HPR Part I funds and the report after acceptance by the JHRP Board will be submitted to the ISHC and the BPR for their review, comment and acceptance.

Respectfully submitted,

Harold L. Michael

Harold L. Michael
Associate Director

HLM:maz

cc: F. L. Ashbaucher
W. L. Dolch
W. H. Goetz
W. L. Grecco
G. K. Hallock
M. E. Harr
R. H. Harrell
J. A. Havers
V. E. Harvey
G. A. Leonards
F. B. Mendenhall
R. D. Miles
C. F. Scholer
M. B. Scott
W. T. Spencer
H. R. J. Walsh
K. B. Woods
E. J. Yoder

Progress Report

ACTIVITY - ACCESSIBILITY MODELS OF TRIP GENERATION

by

Tammam Z. Nakkash
Graduate Instructor in Research

Joint Highway Research Project

Project: C-36-69E

File: 3-7-5

Prepared as Part of an Investigation

Conducted by

Joint Highway Research Project
Engineering Experiment Station
Purdue University

in cooperation with the

Indiana State Highway Commission

and the

U.S. Department of Transportation
Federal Highway Administration
Bureau of Public Roads

Purdue University
Lafayette, Indiana
May 6, 1969

ACKNOWLEDGMENTS

The author wishes to express his sincerest appreciation to Professor William L. Grecco for his guidance, assistance, and encouragement throughout the conception and execution of this research. Professor Grecco's critical review of the manuscript was also most valuable.

Gratitude is extended to Professor Harold L. Michael for the opportunity to undertake this work and for his review of the manuscript. The advising of Professor Virgil L. Anderson on the Statistical aspects of the research and his review of the manuscript are very much appreciated. Thanks is also due to Professor James R. Buck for his review of the manuscript.

The author is grateful to the Joint Highway Research Project of Purdue University and the Indiana State Highway Commission for the financial support which, in addition to the assistance of the Bureau of Public Roads of the U. S. Department of Transportation, made this research possible.

Acknowledgment is also due to personnel of Barton-Aschman Associates for their furnishing of data and assisting in its preparation for analysis. Mr. Gary M. Jouris is thanked for his counsel on statistical topics.

The author wishes to extend his appreciation to many fellow graduate students whose interest, suggestions, and encouragement were valuable for this research.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

ABSTRACT

Nakkash, Tammam Zaki. Ph.D., Purdue University, June, 1969. ACTIVITY-ACCESSIBILITY MODELS OF TRIP GENERATION. Major Professor: William L. Grecco.

This research examined the trip generation process with the specific purpose of evaluating the effect of activity-accessibility variables on trip generation. Another aspect of accessibility was studied by stratifying the zones of the study area by location. In contrast to the traditional trip generation procedures where the trips generated by a zone are considered to be a function of the characteristics of only the zone itself, the models proposed by this research take into account also the characteristics of all the other zones and the transportation network connecting them.

Operationally, the accessibility of a zone to each activity was defined as the sum of the products of the size of each activity in each zone of the study area and the friction factor corresponding to the travel time between each zone and the zone under consideration. The friction factors were those derived from calibrating a gravity model of trip interchange for various trip purposes. The accessibility of each zone was then expressed as a percentage of total accessibilities of all zones of the area and referred

to as relative accessibility. Criteria were also set to differentiate the zones of the study area into two strata: central and non central areas.

Utilizing data from the 1964 surveys for a comprehensive transportation study in Indianapolis, Indiana, four sets of trip generation regression equations were developed for each of thirteen trip purposes. One of the developed sets was a recalibration of the Indianapolis Regional Transportation and Development Study (IRTADS) trip generation equations. This set was developed by the traditional procedures, that is, the independent variables were limited to demographic, socio-economic, and land use variables. A second set was developed in which relative accessibility variables were included among the independent variables. Two more sets were developed, each corresponding to one of the former two sets but calibrated with data stratified according to the zone's location in the central or non central area.

The developed models were compared as to their statistical strength. Considering the factor of location always improved the statistical strength of the trip generation models. Including relative accessibility variables improved the statistical strength of trip attraction models more than that of trip production models. Stratification alone improved the models more than including relative accessibility variables only.

1985 forecasts of the demographic, socio-economic, and land use variables together with estimates of travel time on the proposed future network were inputs to the four sets of developed models to forecast trip productions and attractions. Comparison of these forecasts indicated that stratified models forecasted, on the average, more trips for zones of the non central area and less for zones of the central area compared to forecasts made by unstratified models. Forecasts by models with accessibility variables and stratification were significantly different, on the average, than forecasts by basic IRTADS models. There were some trip purposes for which IRTADS models indicated, on the average, larger trip forecasts than the accessibility models with stratification. Further analysis on those trip purposes showed that the zones where the IRTADS models did not overforecast trips were in the vicinity and along corridors defined by the major thoroughfares of the study area.

The results of this research recommend that future relative accessibility of zones of the study area be considered in estimating future trips. The process would be iterative and would be terminated when equilibrium is reached between the forecasted demand for transportation (future trips) and planned for supply of transportation (future network).

CHAPTER I. INTRODUCTION

In recent years one has witnessed a great increase in urban transportation planning activities in many American cities. This increased activity was in response to the challenging urban transportation problem, one of the major urban problems of contemporary cities. The safe and efficient intracity movement of goods and people is very essential for the social, cultural, and economic health of an urban area.

Urban Transportation Planning

The urban transportation problem is the product of many interacting factors. The enormous population growth in urban areas and their expanding areal extent as a result of the redistribution of population, the improved standard of living due to increased affluency, and the subsequent greater reliance on private automobiles are only some of those factors. Together with those size-related features of the problem, the temporal aspects induce periodic high demands for transportation. This, of course, is due to the interdependency of human activities which simultaneously occur, essentially, during the eight-hour work-day; starting and ending at rather definite times.

A recognition of the immense complexity and the vast dimensionality of the urban transportation problem is a prerequisite to any attempt at solving the problem.

The symptoms of the urban transportation problem were first most apparent in the congestion of the streets; thus early attempts at reducing the severity of the situation were traffic volume oriented. The failure of projections based on traffic volumes' factoring to cope with the need for accurate forecasts, to be used in highway planning and design, has directed the effort of transportation planners since the 1940's toward more sophisticated approaches. The development of Urban Transportation Planning (UTP) was a step in this direction.

UTP views the traffic condition of the city streets as an element within the framework of the larger urban environment. Projections for transportation planning include in addition to the expected amount of travel its temporal and locational pattern; and consider all transportation modes. UTP is a continuing process, projections and plans are periodically adjusted to conform to changes in the complex picture of the urban environment.

UTP has developed into an intricate process which relies heavily on surveys. An enormous body of data is collected, usually, on a sampling basis. The data include information on the demographic, social, and economic characteristics of the area. An exhaustive inventory of

the existing transportation facilities and their characteristics is also conducted. The land use survey supplies the information on the quantity, intensity, and location of land classified by use. Detailed floor area information are collected for commercial, services, and educational uses.

The backbone of the urban transportation planning surveys is the Origin-Destination (O-D) Study. This study consists of a home-interview, a truck-taxi survey, and a cordon count. The results of this study establish the travel patterns in the base year and provide the data inputs for model formulation.

The socio-economic, demographic, and land use information are, usually, summarized by zone. The zone is, usually, the smallest areal unit considered in all further analyses and projections.

The transportation network is coded into a topological network of links and nodes. Access to, and egress out of the network is accomplished via zone centroids. The topological network carries all the information of the physical transportation network relevant to the planning process. This abstract numerical representation of the road and transit network is amenable to mathematical manipulation.

The travel information obtained from the home-interview, and truck-taxi survey are utilized to explain trip generation of the zones of the study area. The socio-economic

and land use characteristics of the zone are considered to be the explaining variables.

A model of zonal trip interchange is calibrated utilizing the characteristics of the transportation network and the present travel pattern. The parameters of the calibrated model reflect the propensity of each zone to interchange trips with the other zones, and the rate of decay of this trip interchange with increased zone-to-zone separation, measured in travel cost, travel time, or travel distance. These relationships are assumed either to stay the same for the future, or to change in a predictable manner.

The distribution of trips is then assigned to the transportation network yielding link volumes. A successful assignment which reasonably matches ground counts in the base year indicates that a similar assignment of forecast trip interchanges should be valuable in indicating deficiencies.

Socio-economic, demographic, and land use data are then, separately, forecasted for the future design year. This provides the basis for a future trip generation forecast. Future trip generation is an input to estimate future zonal interchanges using the calibrated distribution model. Future trip assignment is the basis for identifying possible future deficiencies, evaluating proposed networks, and proposing alternative plans. Figure 1 shows the major steps in UTP.

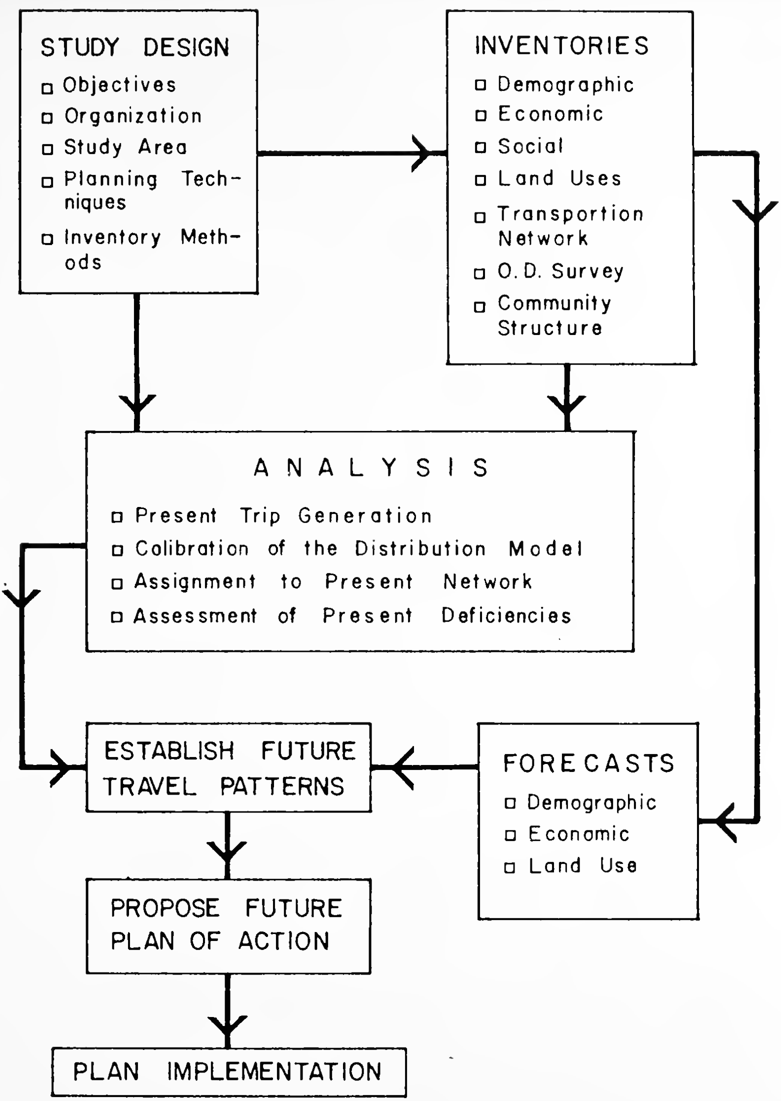


FIGURE I – URBAN TRANSPORTATION PLANNING

Trip Generation in the Urban Transportation Planning Process

Apart from the socio-economic, demographic, and land use forecasts trip generation constitutes the first step toward establishing the future travel pattern. The accuracy of the future trip distribution in forecasting design year trip interchange cannot be any better than the accuracy of the trip generation forecasts, except due to chance.

The ultimate purpose of the trip generation analysis is to arrive at an estimate of the trip ends generated at each analysis unit of the study area. Trip generation techniques try to establish a relationship between the demographic and socio-economic characteristics of the population of an analysis unit and its trip generation. Similarly, the intensity, character, and location of different land uses are related to trip making of the analysis units. These procedures are based on the hypothesis of a causal relationship between population characteristics, land use, and the trip making behavior of people.

Traditionally, trip generation forecasts are established independently of any direct consideration to the transportation network. This, of course, assumes that trips produced at, or attracted to a zone are a function only of the attributes of the zone itself; and are not directly a function of the transportation network which provides the roadway for the trips. Apart from some consideration of the future network in forecasting land uses, this ignores the feedback and interaction between some significant

components of the complex system of the urban environment. A flow chart of the traditional trip generation process is presented in Figure 2.

Purpose and Scope

The purpose of this research was to study the trip generation process, with the specific aim of investigating the feedback from the transportation system on the rate of trip making. Conceptually, there is no strong basis for assuming that trip making is independent of the transportation system. On the contrary, it seems that trips produced by, or attracted to a zone should be a function of the relative accessibility of the zone to different land uses, in addition to the characteristics of the zone itself.

Trip making is a product of the desire for human interaction. Within present day technology, way-of-life, and the requirement of compatible land uses, different daily activities have to be performed at different locales. The possibilities and merits of substituting travel by improved communications are fully recognized [1].* However, it is not anticipated that such substitution will drastically modify travel needs and habits, in our cities, within the next few decades.

Basically, the rate of trip making is a function of two categories of variables. One category represents

---

* Numbers in brackets refer to numbered items in the List of References.

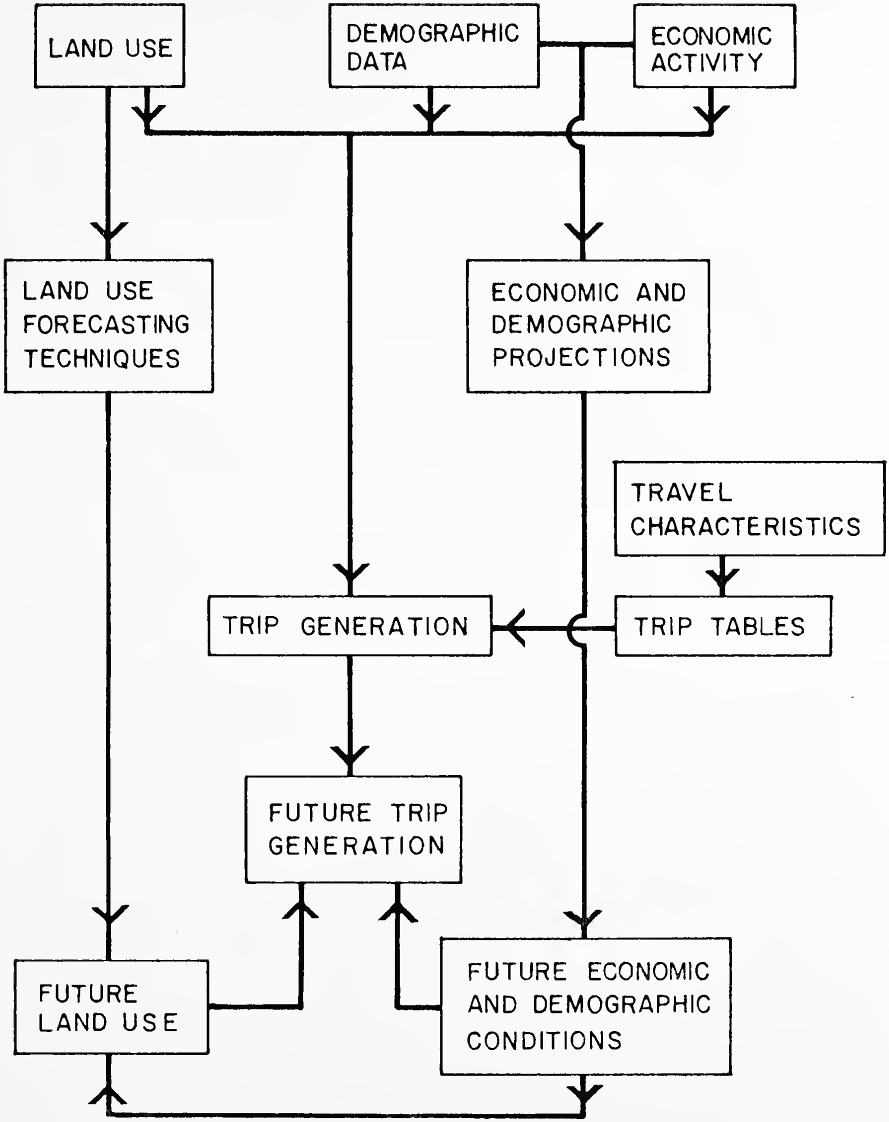


FIGURE 2 — THE TRADITIONAL TRIP GENERATION PROCESS

variables which tend to increase the potential of trip making; the other constitutes the restrictive forces. Variables such as the availability of the vehicle to the residents of a zone, and the percent participation of the residents in the labor force; or the number employed in a zone, and the amount of floor area of different land uses in it are examples of the first category of variables. They measure respectively the potential of trip production or trip attraction. The penalties incurred by travel measured in cost, travel time, or travel distance represent the variable which belong to the second category, the restrictive forces.

This study utilized data obtained from the surveys for the Indianapolis Regional Transportation and Development Study (IRTADS). Multiple linear regression predictive models of person-trip productions and attractions, by purpose, were developed. The developed models differ from the traditional trip generation models. The independent variables were not restricted to socio-economic and land use measures of the zones, but included also measures of the relative accessibility of the zone to different activities and land uses.

The locational aspects that affect trip generation were also investigated. It was hypothesized that central locations in the study area, generally, afford greater accessibility; and the convergence of the street network on the city center favors the core location. The zones of the study area were stratified into two groups: a central and

a non central. This stratification was entered as an independent dummy-variable in the trip generation analysis.

A comparison of the forecast trip generation by the suggested approach with the forecast by the traditional approach was conducted.

CHAPTER II. REVIEW OF LITERATURE

General

The ultimate purpose of the trip generation analysis is to establish a reasonable estimate of the trips produced at, or attracted to some subunit of the urban area, for the design year. This problem has been approached differently by various transportation studies. Due to the difficulty of projecting trips directly, it has been general practice to establish some relationship between trips and other attributes of the urban area and to use this relationship to provide a forecast.

The three main approaches that have been used are: (1) trip rate analysis, (2) cross-classification analysis, and (3) multiple regression. These three approaches and their variations will be discussed later separately.

It is significant to observe that, as yet, "The Model" of trip generation is not available. That is, there is no simple model or set of models which have well defined structure and mathematical formulation that is applicable to any study area when calibrated with data obtained from it. Apparently, a valid conceptual basis for trip generation has not yet been reached.

Trip generation analysis is usually conducted with reference to some subunit of the study area. In this case, and since the O-D study furnishes information on the dwelling unit level, some aggregation of data has to be done. A valid model of trip generation should predict the same total number of generated trips, regardless of the way the study area is subdivided, or the level of data aggregation [2]. The subdivision of the study area into subunits, such as traffic zones, could be done in an infinite number of ways. Urban geographers have recognized this fact

> ...any set of areal boundaries is essentially arbitrary... When one computes the density of an areal unit, he obtains a figure which, though well defined, is necessarily relative to areal boundaries and hence fictitious in a sense [3, p. 37].

When dividing the study area into zones, the interest will be in the among-zones relationship, rather than the variation within the zone. The assumption of homogeneity within the zone is of questionable validity. Supposedly, the smaller the zone, the better the chances that this assumption is valid; but rather the location of the boundary, not only the size of the zone, will affect its homogeneity [3].

Trip Generation Analysis Procedures

The procedures used in trip generation analysis attempt to establish a relationship between the amount of trips generated per zone, and the various demographic, economic and land use characteristics of the zone. They differ only, in

the manner in which the relationship is established.

Land Area Trip Rate Analysis

The method utilizes the numerical relationship between land use and trips which it generates [4]. The rate of trips generated by land areas of different character, location, and intensity are established. Character of the land area denotes its functional use, location refers to the spatial arrangement within the urban area, and intensity of use is usually measured by density.

Trip ends data, obtained from the O-D study, are related to land use information to establish the trip rate per acre for each different land use. The procedure was first used and documented by the Chicago Area Transportation Study (CATS), after it was originally introduced by the San Juan Study [5]. The experience of CATS with this procedure indicated that a forecast of the trips by applying the trip rates of the base year to the land use forecast of the future year resulted in an underestimate. This discrepancy resulted because the trip rate, itself, is a function of other variables which were not considered. The low estimate had to be adjusted by a control total established by a second projection based on car ownership and residential density [6]. An example of trip rates by land area for different land uses is presented in Table 1.

Table 1. Person and Vehicle Trips Generated by Various Land Uses.

Ring	Average Distance From Loop in Miles	Person Trips Per Acre						Vehicle Trips Per Acre (Trucks Weighted)					
		Residential	Manufacturing	Transportation	Commercial	Public Buildings	Public Open Space	Residential	Manufacturing	Transportation	Commercial	Public Buildings	Public Open Space
0	0.0	2,228.5	3,544.7	273.1	2,132.2	2,013.8	98.5	1,336.9	1,081.1	103.4	728.1	461.0	62.3
1	1.5	224.2	243.2	36.9	188.7	255.5	28.8	93.3	162.5	54.6	194.0	116.5	26.2
2	3.5	127.3	80.0	15.9	122.1	123.5	26.5	54.0	64.2	30.6	116.7	50.9	22.7
3	5.5	106.2	86.9	10.8	143.3	100.7	27.8	49.5	66.0	14.7	132.1	46.4	17.5
4	8.5	68.3	50.9	12.8	212.4	77.7	13.5	35.5	43.8	12.4	165.4	33.8	11.6
5	12.5	43.0	26.8	5.8	178.7	58.1	6.1	25.3	23.4	7.1	150.2	29.3	4.4
6	16.0	31.2	15.7	2.6	132.5	46.6	2.5	19.4	14.7	2.9	111.7	24.2	1.8
7	24.0	21.2	18.2	6.4	131.9	14.4	1.5	13.3	15.7	6.4	115.3	7.2	1.0
Study Area Average		48.5	49.4	8.6	181.4	52.8	4.2	26.1	38.6	10.2	144.6	24.2	3.1

Source: Chicago Area Transportation Study (CATS), Reference [4].

Cross-Classification Analysis

Cross-classification is a technique applicable when large numbers of observations are available. The technique requires data to be summarized in cross-tabulations, so that the frequencies or the average values of a dependent variable can be obtained for different combinations of values, by classes or ranges, of two or more independent variables [7]. This type of analysis does not rely heavily on any assumptions, about the distribution of the variables. Thus, it could be thought of as a nonparametric technique [8].

Cross-classification has been used as a research tool in the analysis of trip generation. Oi and Shuldiner reported cross-tabulations, such as trips per household at various levels of car-ownership and persons per dwelling unit [9]. Recently, cross-classification is gaining recognition by operational transportation studies [10,11,12].

A modification of the cross-classification analysis was developed at the Puget Sound Regional Transportation Study (PSRTS). Later it was used by the Albuquerque Transportation Study to compare results with a regression procedure forecast of person-trips [12]. For home-based person-trip productions, trips per household were calculated. Analysis zones of similar household and enviromental characteristics were grouped together, and average trip rates per household were calculated for each group of zones. In

selecting the independent variables, results of previous research, and variability of the data were the criteria.

The zones were ranked according to the dependent variable from high to low, similarly, they were ranked according to each of the independent variables. A Spearman's rank correlation coefficient rho, of the ranking based on the dependent variable as compared to each of the rankings according to each of the independent variables was calculated.

Average person-trips per household were cross-tabulated at three levels of intensity of environmental characteristics and six levels of household characteristics. The 18-cell table was referred to as the rank-classification matrix, and is shown in Figure 3.

Projections of the independent variables for the forecast year were developed at the zone level. In order to establish a forecast of the trips produced at any zone, one must use the rate of trips per household in the cell of the rank-classification matrix corresponding to the forecasted environmental and household characteristics of the zone. An example is shown in Figure 4.

A similar approach was used to project trip attractions. It is significant that some zones had a small employment base, therefore, in order to insure statistical stability of employment trip generation rates "...zones for which the trip information indicated less than a selected level were omitted from the analysis for determining the groupings of zones." [12, p. 95].

HOUSEHOLD CHARACTERISTICS	ENVIRONMENTAL CHARACTERISTICS		
	Low	Medium	High
Low-Low	(1)*	(2)	(3)
High-Low	(4)	(5)	(6)
Low-Medium	(7)	(8)	(9)
High-Medium	(10)	(11)	(12)
Low-High	(13)	(14)	(15)
High-High	(16)	(17)	(18)

(CELL NUMBER)*

SOURCE: REFERENCE NUMBER 12

FIGURE 3 — RANK CLASSIFICATION MATRIX

HOUSEHOLD CHARACTERISTICS	ENVIRONMENTAL CHARACTERISTICS		
	Low	Medium	High
Low-Low	—	2.88	3.19
High-Low	5.51	6.03	5.29
Low-Medium	7.57	6.94	6.26
High-Medium	7.96	7.53	6.84
Low-High	8.47	8.38	7.79
High-High	9.08	9.54	—

SOURCE: REFERENCE NUMBER 12

FIGURE 4 - RANK CLASSIFICATION MATRIX — AVERAGE PERSON-TRIPS PER HOUSEHOLD

Cross-classification is simple, does not rely heavily on assumptions underlying the distribution of the variables, keeps the study staff closer to the data; but it has, however, serious disadvantages. Large number of observations are needed for conclusive results. The number of observations needed increases with the increase in the dimensionality of the classification matrix. Major shortcomings of this technique were stated as follows by Ezekiel and Fox

> ...it provides no measure of how important the relation shown is as a cause of variation in the factor being studied, or of how closely that factor may be estimated from the others on the basis of the relations shown...[cross-classification]...does not determine the relationships where many variables are involved so satisfactorily as does multiple regression [7, p. 394].

Multiple Regression Analysis

General Multiple regression is a statistical technique to fit a response surface satisfying the least squares criterion. The increased use of multiple regression could probably be attributed to the recent availability of computer codes that can handle the estimation of regression parameters efficiently.

Generally the functional relationship between the dependent and independent variables could be represented mathematically by

$$Y = f(X_1, \ldots, X_p | B_1, \ldots, B_q)$$

where

Y = the dependent (or response) variable

X_i = the ith independent variable ($i=1, \ldots, p$)

B_j = the jth parameter in the function ($j=1, \ldots, q$)

and f stands for the assumed form of the function [13].

In order to make probability statements on the parameter estimates three basic assumptions have to be satisfied:

1. the independent variables are fixed and measured without error;
2. for each given value of the independent variables, the values of the dependent variable are normally and independently distributed; and
3. the variances of the distributions of the dependent variable are all the same [13].

In a trip generation analysis based on multiple regression the previous assumptions are likely to be violated. Statistical tests are available to check if the second or third assumption is violated. Some transformations of the data could be used to correct for a violated assumption. The topic of regression is well documented in the literature on statistics [7,13,14].

Linear Versus Nonlinear Regression. The distinction between linear and nonlinear regression is in the functional form of the model. Linear regression, as the name indicates, refers to regression models which are linear in the

parameters. Some models with exponential parameters are intrinsically linear since they can be put in linear form by a transformation [14]. The estimation of the parameters of a linear model is reached by solving a set of simultaneous linear equations, referred to as the Normal Equations. The estimation of the parameters of a nonlinear model varies in complexity with the form of the model. Search techniques are usually utilized for this purpose.

Linear regression has been more popular than nonlinear regression in trip generation analysis. Linear regression is simpler, and does not require _a priori_ knowledge of the exact functional form of the model. One of the earliest documentations of the use of linear regression was by Wynn [15]. Nonlinear regression, however, has the advantage of forcing the researcher to look at the relationship between the variables. Nonlinear regression was used to estimate recreational trip rates in research reported by Matthias [16].

Constrained Regression Estimation. The assumptions of regression state that the independent variables are fixed and uncorrelated. The independent or explanatory variables in a regression analysis of trip generation, in general, violate these assumptions. In the extreme case, where a very high correlation exists between some of the independent variables, the computational procedure fails, since this implies matrix singularity and the Normal Equations do not

have a unique solution. If this occurs, enough warning will be received. A less-than-perfect correlation between some of the independent variables will result in dividing their effect on the dependent variable rather arbitrarily. Many times this yields conceptually invalid parameter estimates (regression coefficients). If theory can provide some bounds on the values of some of the coefficients of the independent variables, then it seems appropriate to use this information in developing the regression models; thus avoiding structural mis-specification. Finding the parameters' estimates of a constrained regression model is a quadratic programming problem. Meyer and Glauber provide the basic mathematical formulation for this problem [17, pp. 184-196]. Recently reported research utilized this technique in estimating urban passenger travel behavior [2].

The Uses of Dummy-Variables in Regression. Some of the independent variables that are relevant to trip generation are impossible to measure on a continuous scale. Cross-classification is one of the useful techniques in such a case, however, it has the disadvantages stated earlier. The dummy-variables technique handles this case in the framework of multiple regression. In the context of trip generaation household characteristics could be stratified with respect to various factors, with different levels for each factor. Each household characteristic will have a dummy-variable associated with it, except one class to avoid

having more coefficients to estimate than independent Normal Equations. The dummy-variable associated with each class will have a value of zero except for the one class to which the observation belongs will have a value of one. The earliest reference in the literature of statistics on the use of dummy-variables is, probably, the article by Suits [18].

The use of dummy-variables in trip generation studies is gaining acceptance and is being suggested by many of the recent publications [8,19,20]. Its use was reported in a research paper by Fleet, Stowers, and Swerdloff [21].

Site Analysis

Data obtained from the O-D survey is of questionable applicability in predicting nonresidential trip generation, especially if the purpose of the trip estimation is system design rather than area-wide planning [22]. While steady growth could be accounted for on the zone level, the emergence of a major trip generator would probably best be treated separately. Data for trip generation analysis of major generators should probably be collected at the site (generator), rather than at the dwelling units.

In the base year analysis, individual site investigation can supplement the zone-level analysis. In the continuing phase, periodic checks at the site can provide continuous updating of the trip generation. Site analysis has not yet been used extensively in operational transportation

studies. Significant work has been reported by the California Division of Highways [23,24]. Special traffic counts are done at selected sites for a period ranging from one day to seven days, depending on the situation. For the seven-day or five-day (Monday through Friday) period, trip averages per 24 hours are related to some of the characteristics of the site such as floor area, number of employees, number of patients, or student enrollment for generators such as shopping centers, offices, hospitals, or schools respectively [25].

Both, interview techniques and volume counts, were reported to have been used by the Chicago Area Transportation Study to analyze trip generation at a high rise apartment building and at O'Hare International Airport [26].

Factors in Trip Generation

Urbanization is an expression of the interdependency of human activities. The concentration of people in urban areas results from the pursuit of opportunities for interaction. Face-to-face interaction and person-to-place contact presume spatial proximity. Urban travel is done, in general, for these purposes. Excluding any possible utility of the mere act of travel, the demand for transportation is a derived demand. It is not sought for its own sake, but rather to satisfy some desire for interaction.

The Trip Generation Mechanism

For the proper understanding of the generation of trips, and for predicting future trip generation, travel should be observed within the framework of the total urban picture. Stated otherwise, trip generation is better analyzed and described by the same scales which measure the complicated activities of the urban area.

Travel is best and most completely described, as suggested by Mitchell and Rapkin, by the following characteristics: (1) volume, (2) composition, (3) distance, (4) time rhythms, (5) location, (6) density, (7) kind of trip, (8) kind of establishments at origin and destination, and (9) trip maker characteristics [27]. Among these characteristics, the last three are, probably, the most relevant to trip generation.

Trips could be classified according to purpose such as: work, shopping, school, recreation, or personal business. For the purpose of the distribution models trips are usually classified also into production trips and attraction trips, and also as home-based or non home-based trips. Person-trips could also be classified according to the mode of travel such as walking-trips, auto-trips, or transit-trips. Most O-D studies ignore walking-trips which, of course, introduces a bias in the calibration of any person-trips generation model.

In establishing a method for identifying independent variables, three methods were suggested by Mitchell and

Rapkin:

1. systems based on kind of trip, that is, trip purpose;
2. systems related to kind of establishment or area, that is, land use, either by parcel or by area; and
3. systems based on process of action or roles [27].

A comparable set of methods reported by Shuldiner include:

1. land use;
2. activity-purpose; and
3. others [22].

Mitchell and Rapkin suggest that travel is composed of dispersive, assembling, and random components [27]. The trips generated at the home are considered dispersive, they might be predicted by household characteristics. The destination end of these same trips could be predicted by the land use at each of the destinations. In the prediction of trips, two types of variables are available, causative and symptomatic [28]. For example, the school enrollment in a zone is a causative variable of the school trip attractions. Also, the number of single family dwellings in the same zone could be correlated with school-trip atractions. The number of single family dwellings do not cause the school-trip attractions, rather it is a symptomatic variable [28]. Since most of the characteristics of the urban area are related, relationships based on causative variables would be

more stable with time than those based on symptomatic variables; thus providing a better predictive model.

Factors Used to Estimate Residential Trip Generation

Residential trips are of the dispersive type. They are best estimated by relating them to the characteristics of the household.

Household Size. Oi and Shuldiner stress the effect of family size on trip generation [9, p. 85]. Stowers and Kanwit reported a higher standardized regression coefficient (Beta coefficient) in correlating trips to family size than to automobiles owned [20, p. 50]. Williams and Robertson also found a high effect of family size on trips [29].

Car Ownership. Oi and Shuldiner state that car ownership is the one variable which exhibits the closest association with reported trip generation rates [9, p. 86]. They further state that car ownership depends on other variables such as real income, price, and financing terms, as supported by studies of the automobile market [9, p. 87]. The impact of car ownership on trip rates is twofold. It first increases the proportion of the residents who are trip makers. Second, increased car ownership enhances the intensity of travel [9].

Janes reported that the average number of trips per vehicle for two-vehicle households is almost one-third less per vehicle, than for one-vehicle households [30, p. 21].

That is the increase in vehicles per household does not proportionately increase the trips generated by these households.

Income. Oi and Shuldiner suggest deleting income from the analysis, because it is highly correlated to many factors that are more highly correlated to trips [9, p. 104]. Stowers and Kanwit, however, rank income third after family size and car ownership in the order of significance relative to trip making [20, p. 50].

In his study of the effects of car ownership, Wynn divided the population into four income classes and compared 1948 to 1955 data and found the following:

1. the number of cars for the highest three income brackets was almost constant, and
2. between 1948 and 1955 the cars per resident was approaching a constant for all income classes.

He thus concluded that "car ownership is much less significant than level of income." [15, p. 26].

Occupation of Head of Household, and Social Status. In a study by Walker, rate of trips generated by a household were found to be a function of the occupation of the head of household, if the population per household and car ownership were controlled [31]. Households were occupationally stratified by "white collar" and "blue collar" employment, or as high, medium, and low occupational groupings.

Earlier, Oi and Shuldiner tested the hypothesis of the significance of occupation of head of household on the rate of trips generated per dwelling unit by a covariance model. Although the test was significant, they questioned the validity of the linearity assumptions used [9, p. 116].

In their attempt to study the effect of social status, Oi and Shuldiner investigated the significance of the social attributes of an area described by other indices: (1) social rank, (2) degree of urbanization, and (3) extent of segregation. These social area indices are referred to as the Shevky-Bell Typology [32].

Another social factor proposed by Stowers and Kanwit is the "stage in the family life cycle" [20, p. 50]. It ranked fourth after family size, automobiles owned, and income in explaining the variability in household travel habits.

The Distance from the Central Business District (CBD). It has been suggested by Wynn that the distance from the CBD is an important factor in trip generation rates [15]. Some researchers have argued, however, that the distance from the CBD could be a proxy variable. The distance of a dwelling unit from the CBD is, in fact, very much related to many household and environmental characteristics, such as family size, automobiles owned, income, stage in the family life cycle, and residential density of the neighborhood. This idea was illustrated by Mertz et al. [33]. Oi and

Shuldiner have shown that the importance of this factor is insignificant when the correlation of trip generation rates with car ownership and household size is accounted for [9]. A similar conclusion was reached by Stowers and Kanwit [20].

The Type of Dwelling Unit. The type of a dwelling unit reflects the style of life of its occupants; and it is, probably, related to its distance from the CBD. In studying the relationship between dwelling unit type and rate of trip generation, Oi and Shuldiner conclude that the effect of the relationship is not nearly so pronounced as that of income or occupation [9, p. 121].

Additional Remarks. The previous discussion on the factors affecting residential trip generation was oriented toward a dwelling unit analysis. When the analysis is at the zone level, appropriate equivalent variables are usually used. Population, labor force, and school-age population are some examples. Such demographic variables are correlated with residential trip generation at the zone level. Population was noted to be highly correlated with residential trip generation by Sharpe et al. [34]. It should also be noted that measures of the intensity of the use of land are only possible to calculate in areal units of analysis, such as the zone or district.

Factors Used to Estimate Nonresidential Trip Generation

In developing models for estimating nonresidential trip generation two basic approaches are in use: (1) Land-Use Based Models, and (2) Activity-Purpose Models [22, p. 74]. The basis for the above classifications is to facilitate the review of each model and does not necessarily reflect basic differences between them.

Factors Related to Land Use. A basic concept in transportation planning is the recognition of the functional relationship between travel patterns and land use, which was extensively analyzed by Mitchell and Rapkin [27]. Land-use based models establish relationships between the amount, location, and intensity of use of different classifications of the use of land and the amount of trips generated by them.

In providing the basis for such a procedure, a generalized land use classification system is needed. The same system is not used by all urban area transportation studies. Even within the same urban area, variations in trip rates within generalized land use classes have been far too large [22, p. 74]. Another source of difference between urban transportation studies is introduced by the methods by which trip rates are calculated. Trips per acre, and trips per thousand square feet of floor area are not comparable because of the range in establishment sizes, building heights, and land coverage. In addition, gross acre is

sometimes used, while on other occasions net acre is used.

The use of the floor area, as a basis for the analysis, instead of the acreage, has the advantage that floor area is a better measure of the intensity of use. It reflects the effect of the multi-story structures, and thus the character of the area also. The limited use of floor area measures can be attributed to the difficulty and expense of obtaining those measures.

Black compared land area versus floor area as a predictor of nonresidential trip generation. He concluded: "Contrary to expectation, floor area does not seem consistently better than land area. Further more, floor area trip rates are not uniform throughout a metropolitan area, but increase as the density decreases." [35, p. 1].

Factors Related to Land-Use Activity. This approach recognizes that the activity which takes place on land is, probably, more related to trip generation than the land area in different uses. Variables such as employment and its various subclassifications, retail sales, and school enrollment are examples of activity measures. Land-use activity measures are good predictors of trip generation by purpose.

Black compared three parameters of nonresidential trip generation: land area, floor area, and employment. He used data collected by CATS. His findings indicate that no one particular measure is best for predicting trip generation for all land use categories. He concluded: "Floor area

seems best for commercial, employment for manufacturing, and land area for public buildings, public open space, and transportation." [35, p. 1].

Accessibility Considerations in Trip Generation in Previous Investigations

Little has been done, as yet, on accounting for the effects of the changes in the transportation network on the trip generation in an operational transportation study. Almost in all the cases, the network consideration were within the framework of the competitive models of trip distribution.

Fort Worth Study

The Fort Worth Study attempted to allocate the total number of trips of a certain purpose which were derived from estimates of household trip productions from all zones. This was used in lieu of estimating the trips by purpose attracted to each zone. Shuldiner comments that this approach is most suited for estimating highly competitive trips, such as shopping trips [22]. The following presentation of the method used by Fort Worth Study relys heavily on Shuldiner's article [22]. A set of attraction indices, which are weights that assign to each land use factor its relative importance in the trip attraction model was developed. Considering the home-based other trip purpose category, for example:

Let

Y_i = the basic attractiveness of zone i for home-based other trips relative to all other zones. Basic attractiveness is used here to designate the attractiveness of zone i without regard to its location or accessibility relative to all other zones;

P_i = population in zone i,

C_i = commercial employment in zone i,

I_i = industrial employment in zone i, and

O_i = other employment in zone i.

With the variables P_i, C_i, I_i, and O_i is associated a set of attraction indices: AI_P, AI_C, AI_I, and AI_O respectively, which are weights that assign to each variable its relative importance in the trip attraction model. Then, Y_i = $AI_P(P_i) + AI_C(C_i) + AI_I(I_i) + AI_O(O_i)$.

"The Y_i's are then combined with appropriate friction factors to distribute the total number of home-based other trips among the various zones by means of the gravity model." [22, pp. 81-82]. This is, actually, a variation of the classical gravity model of distribution; but, instead of using trip ends as the measure of attractiveness of the zone the Y_i's, as defined above, were used.

Southeast Connecticut Area Transportation Study (SEATS)

SEATS made extensive use of the relative attractiveness concept [22]. Employment was used as the single measure of

activity. Nonresidential land was classified into one of the eight following categories: (1) industrial, (2) personal service, (3) business service, (4) institutional, (5) recreational, (6) commercial amusements, (7) retail, and (8) other. Accordingly industrial employment in a zone is the measure of industrial activity, institutional employment in the same zone is the measure of institutional activity, and retail employment is a measure of the retail activity in it, etc. This measure of activity was called the destination zone factor. The trips to the various land uses were then grouped into 3 classes: home-based long, home-based short, and non home-based. Trip type factors for each class of trips to each land use activity by the total employment in that activity. The last step was to develop an attraction index for each zone for each class of trip as a function of the destination zone factors and trip type factors.

For a single class of trips, home-based long, the procedure in notational form is as follows:

Let

k = a subscript denoting a particular land-use activity ($k = 1, \ldots, m$); and

i = a subscript denoting a particular zone ($i = 1, \ldots, n$).

Further let

$$F_{Lk} = \frac{L_k}{\sum_{i=1}^{n} A_{ki}} \quad , \text{ where}$$

F_{Lk} = trip type factor for land use k,

L_k = total number of home-based long trips to land use k in all zones, and

A_{ki} = destination zone factor for land-use activity k, in zone i.

And let

$$I_{Li} = \sum_{k=1}^{m} F_{Lk} A_{ki}, \text{ where}$$

I_{Li} = attraction index for home-based long trips to zone i.

The attraction index, computed above, is used to distribute home-based long trips among the various zones. The basic assumption of the model is that the trip type factor for kth land use F_{Lk} is the same for all zones.

Baltimore Metropolitan Area Transportation Study (BMATS)

The relative attractiveness of each traffic zone was related directly to several types of system variables. The study report stated that network variables which indicate the level of transportation service available to residents were difficult to determine, "since in past studies no standard parameters have evolved." [36, p. 96]. The highway accessibility index, as a measure of highway service, was defined as the inverse of the sum of the travel time from each zone to every other zone. Time was measured from the minimum time path trees. The highway accessibility index reflects the standard of service as measured by average speed on the street system [36]. The transit service index

represents the frequency of transit service available to each zone. It was calculated by totaling the number of transit trips accumulated for each route. These two indices were used with other variables to build multiple regression models of trip generation.

Indianapolis Regional Transportation and Development Study (IRTADS)

IRTADS experimented with measures of accessibility that were adaptations of the accessibility indices produced by the gravity model. Instead of using trip ends as a measure of attractiveness for zones, as normally is the case in the gravity model, IRTADS used total population or employment in each zone. Using travel times from both the IRTADS highway and transit networks, the gravity model program was run to produce both highway and transit accessibilities to population or to employment [37]. Only three of the multiple regression equations of trip generation included the accessibility variables, and each with a negative coefficient. Eventually the accessibility indices were dropped from the trip generation analysis [38].

Other Transportation Studies

In an unpublished paper by Corradino on the trip generation of Sharon-Farrel Study Area, the set of independent variables included some measures of the highway network variables [39]. These variables included:

1. reciprocal of time to center of gravity of population;
2. reciprocal of time to center of gravity of employment;
3. $\log_{10}$ of time to center of gravity of population; and
4. $\log_{10}$ of time to center of gravity of employment.

In this study, Corradino states that the travel time variables were used as measures of decentralization [39, p. 7]; rather than as measures of accessibility.

In a tabular summary of trip generation equations developed by several transportation studies and reported in a paper by Bellomo and Schultz [40], the Boston Study equations contained four accessibility variables:

1. accessibility by highway to all jobs;
2. accessibility by transit to all jobs;
3. accessibility by highway for all population; and
4. accessibility by transit for all population.

Unfortunately no further details were available.

CHAPTER III. DEVELOPMENT OF THE ACTIVITY-ACCESSIBILITY CONCEPT

The Interactance Hypothesis and Gravity Concepts of Human Interaction

In very general terms, the gravity concept of human interaction hypothesizes that the force of interaction between two human activity areas is directly proportional to some function of their population, and inversely proportional to a measure of the spatial separation between them. The hypothesis relates the sizes of the two populations to the impedence or friction induced by the intervening space. This could be stated formally

$$I_{ij} = \frac{f(P_i,\ P_j)}{g(D_{ij})}$$

where

I_{ij} = interaction between area i and area j;

P_i, P_j = the population of areas i and j, repsectively;

D_{ij} = a measure of the separation between area i and area j; and

f and g stand for the assumed form of the functional relationship of population and separation respectively.

Early Formulations of the Gravity Concept

The earliest known explicit formulation of the gravity concept of human interaction was, probably, by Carey [41]. He argued that the same fundamental laws govern both physical objects and human behavior. Carey stated:

> Man, the molecule of society, is the subject of Social Science....The great Law of Molecule Gravitation [is] the indispensable condition of the existence of the being known as man....The greater the number collected in a given space, the greater is the attractive force that is there exerted....Gravitation is here, as everywhere, in the direct ratio of the mass, and the inverse one of distance [42, pp. 42-43].

This concept was later used by Ravenstein and Young to explain migration phenomena. Ravenstein presented empirical data as supporting evidence that migratory movement tends to be toward the larger cities, and that the volume of movement decreases with the increase of distance between the origin and destination [43,44]. Young hypothesized that the number of migrants drawn from several sources, to a single destination, varies directly as the "force of attraction" of the destination and inversely with the square of distance between the source and the destination [45].

Reilly's "Law of Retail Gravitation" uses the gravity concept to explain the drawing power of a retail center. In his formulation, he postulated, that a retail center will attract trade from its surrounding area in direct proportion to the population size of the center and inversely to the square of the distance from the center to the

individuals attracted [46]. The same formulation provides a basis for determining the equilibrium point between two retail centers competing for trade. All the early formulations, as could be recognized, were partial; in the sense that the size of only one of the interacting groups was considered.

Structural Formulation of the Gravity Concept

In the 1940's, the concept of human interaction was structured by Stewart and by Zipf following the gravitational analogy [47,48]. The formulation by Stewart was in terms of Newtonian physics, namely that the "force" of interaction between two groups is directly proportional to the product of the population, and inversely proportional to the square of the distance on the line joining them. Stewart defined the "force" mathematically as follows:

$$F_{ij} = \frac{P_i \, P_j}{D_{ij}^2}$$

where,

F_{ij} = the force of interaction between concentrations i and j;

all other terms as previously defined.

Following the analogy from physics, Stewart defines the "energy" of interaction between two points i and j as:

$$E_{ij} = k \frac{P_i \, P_j}{D_{ij}}$$

where k = constant of proportionality, analogous to the gravitational constant of physics.

Stewart extended his analogy from physics to include the concept of population "potential," which measures the intensity of the possibility of interaction. The potential at point i, or the intensity of the possibility of interaction of an individual at i, with respect to a group at point j is

$$V_{ij} = k \frac{P_j}{D_{ij}}$$

where

V_{ij} = potential at i of the population at area j.

The total potential at i, that is, the intensity of the possibilities of interaction between an individual at i and the population at all the other points would be:

$$V_{i.} = k \sum_{j=1}^{n} \frac{P_j}{D_{ij}}$$

In the above formulation, it is important to include the potential of the population of a certain location with respect to an individual at the same location. That is, it includes the term P_j/D_{ii}, where D_{ii} is usually taken as the average distance from the center of location i to its boundaries [41].

Variations in the Distance Function

The function of distance in the potential and gravity formulations has been a source of controversy. Experimental

results reported by Price, Iklé, Carroll, and others support the argument that the function of distance in the basic formulations of population potential is a simple inverse relationship of distance raised to some power other than one [49,50,51]. Based upon empirical studies, the distance exponent ranged from one-half to over three. Anderson has suggested that the exponent, itself, is a variable, inversely proportional to the population size [52]. Consider, for example, two centers of population P_j and P_k, were $P_j > P_k$, both centers are equidistant from a third center P_i, that is $D_{ij} = D_{ik}$; then the potential of the smaller center P_k will be more reduced by the distance than the potential of P_j. That is in general

$$V_{ij} = k \frac{P_i}{D_{ij}^{a}}$$

where,

$$a = f(\frac{1}{P_j}) \quad .$$

Carrothers argues, however, that the exponent is more likely inversely proportional to distance, rather than to population [41]. "That is an extra unit of distance added to a long movement is of less importance than an extra unit added to a short movement." [41, p. 97]. Thus the expression of the exponent of distance is:

$$a = f(\frac{1}{D_{ij}})$$

Price considered in addition to the distance the direction of the migration [49]. He considered migration movements as vectors having length and direction, the length corresponding to the distance of migration, and the direction depending on points of origin and destination.

Variations in the Population Function

In studying the interaction phenomenon in different regions, Stewart observed that the effect of population is not always the same. He concluded that there are population characteristics that could make one population group responsible for a higher interaction potential than another. To account for this differential, he suggested using relative weights to account for the population characteristics. Going back to the analogy from physics, this is in fact equivalent to accounting for the difference in the molecular weights of two particles. Stewart, thus, suggests the following formulation:

$$E_{ij} = \frac{k\, m_i\, P_i \cdot m_j\, P_j}{D_{ij}}$$

where

m_i = molecular weight of an individual in i; and

m_j = molecular weight of an individual in j.

Dodd in his formulation of the interactance hypothesis introduces to the equation, factors other than population and distance [53]. These factors referred to as Specific Indices of Influence Level should account for different

characteristics of the population, such as sex, education, or income. The basic energy equation thus becomes:

$$E_{ij} = k\frac{(\Sigma r_i)\quad P_i \quad (\Sigma s_j) \quad P_j}{D_{ij}}$$

where

Σr_i = specific indices of influence level for population P_i, and

Σs_j = specific indices of influence level for population P_j.

Anderson suggested raising the numerator of the basic equation to some power [52]. Carrothers suggested raising the individual population elements in the basic equation to a power, which need not be identical for both populations [41]. When all the variations to the basic formulation of the population and distance functions are considered, Carrothers suggests the following general expression:

$$E_{ij} = k\,\frac{(\Sigma r_i)\ P_i^{\,b}\quad (\Sigma s_i)\ P_j^{\,c}}{D_{ij}^{\,a}}$$

a, b, and c are empirically derived exponents.

Variations of the Basic Formulation

Empirical investigations have suggested different interpretations in drawing analogies to man and distance. So far in this discussion, population has been used as the measure of mass. In studying migration, however, variables such as employment, or the income of the regions or subareas tends to be a more significant measure of mass than

population [54]. In studying market potentials, factors such as dollar volume of sales or income could be more significant than population. The choice of the measure of mass is controlled by the nature of the problem at hand, and by available data. The wide range of variables which could be applicable includes employment, school enrollment, income, telephone calls, newspaper distribution, etc.

Similarly, the appropriate measure of distance need not be restricted to straight line distance. Other measures of impedence such as travel time and/or travel cost has been used [55]. Some problems might make it appropriate to consider road distance, fuel consumption, delay time, or even intervening opportunities as a measure of "social distance" [56].

The Gravity Concept and Urban Trip Distribution

Since the mid 1950's, there has been an increased interest in applying the gravity concept to intra-metropolitan movements. Pioneer work by Voorhees advanced the gravity model of trip distribution; it provided a systematic procedure capable of synthesizing zone-to-zone movements, by purpose, of alternative configuration of land use and transportation facilities [55]. The earliest use of the gravity model in an operational study was reported by studies at Hartford, Connecticut; and Baltimore, Maryland [57]. Earlier the interactance hypothesis was tested with O-D data from Charlotte, North Carolina; and St. Louis,

Missouri; by Wynn and Linder [58].

The general formulation of the gravity model in the framework of intra-metropolitan interchange is:

$$T_{ij} = \frac{P_i \; A_j/D_{ij}^b}{A_i/D_{ii}^b + A_j/D_{ij}^b + \ldots + A_n/D_{in}^b}$$

where

T_{ij} = number of trips produced at zone i and attracted to zone j;

P_i = total trips produced by zone i;

A_j = total trips attracted by zone j;

D_{ij} = measurement of the separation between zones i and j, normally expressed in terms of travel time; and

b = an exponent, to be determined empirically.

In this formulation trip productions and attractions are the measure of mass; travel time is the measure of distance, and it is raised to a constant exponent.

Operationally, and for reasons which will be stated later, the following formulation is used:

$$T_{ij} = \frac{P_i \; A_j \; F_{ij}}{A_i \; F_{ii} + A_j \; F_{ij} + \ldots + A_n \; F_{in}}$$

where

T_{ij}, P_i, A_i = same as defined above; and

F_{ij} = a set of "friction factors," which measure the rate of trip decay as a function of interzonal travel time.

The replacement of the inverse function of distance by a set of friction factors introduces computational flexibility. It does not presuppose a constant exponent because the friction factors are a function of travel time and trip purpose.

The Activity-Accessibility Concept in Trip Generation

As stated earlier in CHAPTER II, the rationale of trip generation has been based on the characteristics of the zone: demographic, economic, and land use, with no consideration for the status of the transportation system. This research is based on the concept that trips generated by a zone of the study area are also a function of the status of the transportation subsystem that serves a zone and connects it to the other zones of the study area. The effect of the transportation system on trip generation will be investigated in the light of the relative accessibility of each zone to various urban activities, and the spatial relationship of the different zones to each other. The development of measures of relative accessibility will be discussed fully later. The concern at this stage is to point out instances where researchers have realized the existence of a feedback from the transportation system on the trip generation phenomenon. To date, little has been done to measure those effects.

Mitchell and Rapkin in their discussion of the systems of urban activities state:

> ...it may be said that the <u>amount</u>* and nature of the movement of persons and goods <u>to</u> and <u>from</u> an establishment (or group of establishments) are affected in varying degree by four factors: the nature and size of the establishment, its position and function in the urban organization, the <u>spatial relationships</u> between the establishment and other establishments in the land use pattern, and the <u>existing channels of movement</u> as related to the establishment [27, p. 18].

Martin, Memmott, and Bone in their discussion of the future trip generation characteristics give the following statements:

> ...it was shown that the transportation studies assume trip generation to be independent of the quality of transportation facilities and the resulting accessibility. This assumption is obviously incorrect, as induced or generated traffic has been observed to be an important segment of the total traffic on new facilities [59, pp. 201-202].

Shuldiner, in investigating nonresidential trip generation, points out:

> ...as far as prediction of future trips is concerned, there is the inevitability of change. All present [trip generation] rates are not fixed and immutable. Rather, they are based on a certain structure of land uses and land-use competition which is constantly changing. Agglomeration, competition, <u>and new transportation</u>* all affect the generation of a given attractor [22, p. 75].

Domencich et al. pointed out that urban transportation planning has developed into the, almost separate, processes of trip production, attraction, distribution, assignment, and modal split. This oversimplification of the analysis of a complex system results in the implicit assumption: "that

---

* Underscoring by this author.

the number of trips generated is independent of the performance of the transportation system." [2, p. 66]. They add that this assumption implies that the policies implemented by transportation planners have no effect on the total number of trips made. Then they, conservatively, conclude:

> While it is possible that changes in the transportation system will not affect trip generation, there is no good reason for making this assumption _a priori_ [2, p. 66].

Walker, in an attempt to develop a theory of trip attraction, examined how land was used and for what purpose. He pointed out that the activity occurring on a piece of land suggests to the people which of their needs could be satisfied by interacting with the particular piece of land. He also observed that the size of land puts a limit on the volume of the activity that could be carried on it. Comparing two pieces of land with identical size and function characteristics, Walker concludes: "the attraction of... land will be directly related to the land's accessibility to the total number of possible interactions perceived by the user." [12, p. 98].

Schneider in studying the relationship of trip generation to access started by a modification of the classical formulation of the gravity model of trip distribution. He then imposed the conservation law of equality of productions to attractions, under the simplifying assumption of complete symmetry between origins and destinations. This could be

illustrated in notational form as follows:

$$T_{ij} = P_i \frac{R_j F_{ij}}{\sum_j F_{ij} R_j}$$

where

T_{ij}, P_i, and F_{ij} are the same as defined earlier in the formulation of the gravity model;

R_j = an undefined quantity measuring that which attracts people to a place, rather than A_j as in the standard formulation.

Satisfying the conservation law, and assuming complete symmetry between origins and destinations, we get:

$$\frac{P_i F_{ij} R_j}{\sum_j F_{ij} R_j} = \frac{P_j F_{ji} R_i}{\sum_i F_{ji} R_i}$$

If we further assume the friction factors are symmetrical, that is, $F_{ij} = F_{ji}$, the relation reduces to:

$$\frac{P_i}{R_i \sum_j F_{ij} R_j} = \frac{P_j}{R_j \sum_i F_{ji} R_i}$$

This can hold true for each pair of origin and destination zones only if the two sides of the equation are separately equal to the same constant c, which yields in general:

$$\frac{P}{R} = c\Sigma FR.$$

Stated in words this relationship says that the trip density, at a point, in terms of origins (or destinations) per unit

of R is equal to a quantity which Schneider called the "access integral around the point." [60, p. 165]. Schneider concludes:

> ...the number of trips at a point is proportional to the accessibility of the point and to its attractiveness to people; trip ends appear at a place because people can and want to get there. This a small shift from the customary point of view in which trips are thought of as occurring to satisfy the craving for fulfillment of trip ends [60, p. 166].

The cited references support the concept of this research: that the trip generation of a zone is not only a function of the characteristics of the zone itself, but also is affected by its location in the urban area relative to other zones and its relative accessibility to different land use activities. It is necessary now to make the transition from the conceptual to the operational level.

The Conceptual and Operational Definitions of Accessibility

The term "accessibility," frequently referred to in the literature, is not very well defined neither conceptually nor operationally. Different authors have different understanding and interpretation of it.

Haig understands accessibility as "contact with relatively little friction." [61, p. 38]. Mitchell and Rapkin comment that overcoming space friction involves two types of costs: transportation cost and site rental [27]. Connecting the two costs of rental and of transportation, Haig maintains:

> Rent appears as a charge which appears that the owner of a relatively accessible* site can impose because of the saving in transportation costs which the use of the site make it possible [61, p. 38].

Hansen defines accessibility as:

> ...a measurement of the spatial distribution of activities about a point, adjusted for the ability and desire of people or firms to overcome spatial separation [62, p. 73].

In this definition, accessibility measures "...the intensity of the possibility of interaction..." [62, p. 73]. As such, it is a generalization of the "population potential" concept developed by Stewart [47].

Martin et al. define the accessibility of a site as:

> ...the weighted summed travel resistance between the particular site and all other sites on which there are mutually interacting activities [59, p. 76].

The operational definitions of accessibility are even more diversified than the conceptual definitions. The simplest operational definition is to measure accessibility of a site to another by the inverse of some function of the distance or travel time between the two sites. This definition is completely devoid of any measurement of the characteristics of the two sites of concern. It reflects, only, the degree of spatial separation between them.

Shindler and Ferreri calculated the "highway accessibility" and the "transit accessibility" of a residential zone to a nonresidential zone by multiplying the reciprocal

---

* Underscoring by this author.

of the travel time squared, between the two zones, by the respective total trips attracted to the nonresidential zone by highway or transit. The reciprocal of the travel time squared was referred to as the "friction factor." The "friction factor," according to Shindler and Ferreri, attempts to measure the restraining influence which time has on travel between zones [63]. The implied assumption is that the restraining influence on travel varies as the square of the travel time.

Savigear attempted to measure accessibility quantitatively as a function of travel time and availability of parking. Computationally, accessibility is derived as follows:

Let

n_i = measure of demand on trips from zone i to the zone under consideration, and

t_i = travel time from zone i to the zone under consideration.

Further let

$$\bar{t} = \frac{\sum_i n_i t_i}{\sum_i n_i} \quad ;$$

$\bar{t}$ is a weighted mean average travel time of trips to the zone under consideration from all other zones. An initial estimate of the accessibility, A, of the zone under consideration is: $A = \frac{1}{\bar{t}}$ [64]. Savigear adds that any measure of highway accessibility should take parking into account.

Let

f = probability of being able to park at the zone of destination, for as long as desired.

Then,

A', a better indicator of accessibility is $A' = f.A$ [64].

Savigear's definition, as it could be inferred, is only in terms of travel time weighted by the demand for travel on the system, and the availability of parking at the destination. No considerations are given to the activities at different zones.

A similar operational definition of accessibility was used by the Baltimore Metropolitan Area Transportation Study (BMATS). As a measure of the highway service, the inverse of the sum of the travel times from each zone to every other zone was used and termed the "accessibility index" [36, p. 96]. The travel times were measured from the minimum path trees.

A more comprehensive operational definition of accessibility was suggested by Hansen. It is the ratio of "population over distance" [62, pp. 73-74]:

$$A_{ij} = \frac{S_j}{D_{ij}^a}$$

where

A_{ij} = relative accessibility of zone i to an activity located in zone j;

S_j = size of the activity in zone j (a function of population, employment, land use, etc.);

D_{ij} = travel time or distance between zone i and zone j; and

a = time or distance exponent.

Considering more than two zones:

$$A_{i.} = \sum_{j=1}^{n} \frac{S_j}{D_{ij}^{a}}$$

where

n = total number of zones;

$A_{i.}$ = relative accessibility of zone i to an activity in all zones.

Swerdloff and Stowers suggested that instead of using the travel time or distance raised to an exponent, to use the alternate formulation:

$$A_{i.} = \sum_{j=1}^{n} S_j \; F_{ij}$$

where

F_{ij} = the friction factor, corresponding to a D_{ij} separation as derived from a calibrated gravity model of the area [65].

The hypothesis proposed by this research is that the number of trips generated by a zone are a function of the transportation subsystem that connects the zone under consideration with the other zones of the study area. For this purpose, the term "relative accessibility" will be

defined conceptually and operationally.

It was hypothesized, earlier, that the trips produced by or attracted to a zone are a function of the causal or symptomatic variables modified by the relative ease in overcoming space between that zone and all other zones. Zones with "relatively more accessible" destinations should, in general, produce more trips; similarly, zones that are "relatively more accessible" to origins should, in general, attract more trips. The term "relatively" refers to the zone under consideration as compared to all other zones of the study area. This implies a competitive consideration among zones in generating trips. Zones of similar size activities will attract trips differently according to their locational and accessibility advantages.

As a measure of the ease or difficulty of overcoming space, this study intends to use the set of "friction factors" developed from the calibrated gravity model of trip interchange for the study area. This has the advantage of avoiding the use of a constant exponent of distance, or time. Friction factors as developed by the calibration of the gravity model are a function of travel time and are classified by trip purpose.

This study's definition of "relative accessibility" is a modification of Hansen's [62]. In notational form, "relative accessibility" is computed as follows:

let

i = zone under consideration, $(i = 1, 2, \ldots, n)$;

j = any zone in the study area, including zone i, $(j = 1, 2, \ldots, i, \ldots, n)$;

k = activity under consideration, $(k = 1, 2, \ldots, m)$;

ℓ = trip purpose, $(\ell = 1, 2, \ldots, p)$;

S_{jk} = size of activity k in zone j;

$F_{ij(\ell)}$ = friction factor corresponding to the travel time from zone i to zone j for purpose ℓ; and

$A_{i\cdot k(\ell)}$ = accessibility of zone i to activity k for purpose ℓ.

Then,

$$A_{i\cdot k(\ell)} = \sum_{j=1}^{n} S_{jk} F_{ij(\ell)}, \text{ and}$$

$RA_{i\cdot k(\ell)}$ = relative accessibility of zone i to activity k for purpose ℓ.

where

$$RA_{i\cdot k(\ell)} = \frac{A_{i\cdot k(\ell)}}{\sum_{i=1}^{n} \sum_{j=1}^{n} S_{jk} F_{ij(\ell)}} \cdot 100 = \frac{A_{i\cdot k(\ell)}}{\sum_{i=1}^{n} A_{i\cdot k(\ell)}} \cdot 100$$

The value of the relative accessibility of zone i to activity k for purpose ℓ, as it is possible to infer from its formulation, could be different for a future year if any or all of the following study area parameters change:

1. interzonal travel time, consisting of interzonal driving time, terminal time, and intrazonal time, or

2. size of activity k, in any or all zones of the study area.

The value of the relative accessibility of a zone to the same activity could be different for different trip purposes. The reason being that the friction factors, corresponding to a certain travel time, are usually different for different trip purposes.

Another aspect of accessibility, neither directly related to travel time nor to the size of activity, is considered. The relationship of location to trip generation characteristics is investigated by stratifying the zones of the study area into two disjoint sets: central and non central. The conceptual basis being that "central sites afford maximum accessibility..." [27, p. 108]. The central area is also, more or less, equiaccessible to the various zones of the study area, because of the convergence of the street system on the city center. This stratification introduces a qualitative factor describing the general arrangement of the land uses and the configuration of the street system.

CHAPTER IV. DEVELOPMENT OF ACTIVITY-ACCESSIBILITY TRIP GENERATION MODELS

The following methodology was used in this investigation. Trip generation models which take into account accessibility variables were developed from data from an operational transportation study. These models were then compared with the conventional models developed as part of the transportation study. Both sets of models were used to forecast 1985 trip generation. The two sets of forecasts were compared by testing for any significant differences, on a zone-by-zone basis, between the two forecasts.

Models that take into consideration the stratification of the zones of the study area into central and non central sets were also developed. This stratification was investigated for the models developed by the transportation study and those developed by this investigation. Also investigated were the ranges of the independent variables for the survey year and the forecast year.

Since the main purpose of this investigation was to compare the sets of developed models, with accessibility variables and zones stratified into central and non central sets, to models developed in the traditional approach by an operational transportation study, care was taken to keep any

factors that might disturb the comparison out of the developed models; so that the comparisons would be most valid.

The decision to develop multiple linear regression models of trip generation using data summarized by zone was mainly in the interest of keeping the results of this investigation comparable to those from an operational transportation study.

Data Preparation

The data used in this investigation were the results of the surveys conducted for the Indianapolis Regional Transportation and Development Study (IRTADS). This decision was based mainly on the availability of data and the fact that the IRTADS survey is recent. The study was at a stage where most, if not all, of the analyses were completed, the forecasts established, and the proposed networks evaluated.

Ready-Available Data

The trip generation analysis for IRTADS was completed in October, 1966; and was reported in a Technical Work Paper [38]. All the models developed by IRTADS were for total person-trips (except for truck and taxi equations). The dependent variables were in the form of productions and attractions suitable for distribution by the gravity model. Nineteen trip generation equations were developed by IRTADS. Two of which were for truck and taxi trip ends, another four

were for control totals, and the rest for person-trips productions and attractions by purpose [38].

The Dependent Variables. This investigation was limited to six trip production purposes, five trip attraction purposes, and two control totals (one for all productions and one for all attractions).

Trip generation equations were developed for the following dependent variables:

1. Home-based work person-trip productions.
2. Home-based shop person-trip productions.
3. Home-based school person-trip productions.
4. Home-based other person-trip productions.
5. Non home-based work-oriented person-trip productions.
6. Non home-based non work-oriented person-trip productions.
7. Total person-trip productions.
8. Home-based work person-trip attractions.
9. Home-based shop person-trip attractions.
10. Home-based other person-trip attractions.
11. Non home-based work-oriented person-trip attractions.
12. Non home-based non work-oriented person-trip attractions.
13. Total person-trip attractions.

It was not possible to develop an equation for home-based school person-trip attractions because the key independent variable, school enrollment, was not available.

Socio-Economic and Land Use Variables. A total of 29 socio-economic and land use variables were originally considered by IRTADS in their trip generation analysis; only 15 were, however, eventually retained in the final equations [38]. Moreover, only 10 of these were available for this study for both the survey year, 1964, and the forecast year, 1985. The socio-economic and land use variables, that were available and used by this study as independent variables in developing multiple linear regression models of trip generation, were:

1. Total employment.
2. Retail employment.
3. Service employment.
4. Retail floor area.
5. Educational floor area.
6. Dwelling units.
7. Labor force.
8. Population.
9. Cars.
10. Single-family dwelling units

Appendix A includes an abbreviation key for all the variables used in this investigation.

Generating Accessibility Variables

Different measures of relative accessibility, to be used as independent variables in trip generation regression models, were established. The operational definition of relative accessibility, presented earlier in CHAPTER III, is restated here:

$$RA_{i \cdot k(\ell)} = \frac{\sum_{j=1}^{n} S_{jk} \, F_{ij(\ell)}}{\sum_{i=1}^{n} \sum_{j=1}^{n} S_{jk} \, F_{ij(\ell)}} \quad 100$$

where

$RA_{i \cdot k(\ell)}$ = relative accessibility of zone i to activity k for trip purpose ℓ;

S_{jk} = size of activity k in zone j;

$F_{ij(\ell)}$ = a friction factor corresponding to the total travel time from zone i to zone j for purpose ℓ;

i = zone under consideration, (i = 1, 2, ..., n);

j = any zone in the study area, including zone i, (j = 1, 2, ..., i, ..., n); and

ℓ = trip purpose, (ℓ = 1, 2, ..., p).

Thus the required inputs are the size of various activities in each zone, and skim zone-to-zone friction factor trees.

Size of Activities. The definition of activity was extended for this purpose to include all the ten socio-economic independent variables that were available from the IRTADS surveys. If the information had not been available

from the above source, the size of activities would have been collected from various sources, such as: employment data, from the State Employment Securities Division; floor area information, from the land use survey; and others, from the home-interview survey or census records.

Skim Zone-to-Zone Friction Factor Trees. A set of friction factors for each of six trip purposes were available from the results of the IRTADS calibrated gravity model [66]. Table B1 shows the values of the friction factors for travel times of one to forty minutes for each of the six trip purposes. The procedure of calibrating a gravity model is well documented, and computer programs are available for that purpose [67].

A binary zone-to-zone trees tape was made available for this investigation by IRTADS. This binary tape was updated (intrazonal and terminal times added) and skimmed to give a skim zone-to-zone travel time binary trees using the program PR-130 [68, pp. A-59 - A-64]. Six skim zone-to-zone friction factor trees were built, one for each of six trip purposes. This was achieved using the program MATCH, written for the IBM 7094. Skim zone-to-zone travel time binary trees tape and the six sets of friction factors were inputs to MATCH. The outputs were on tape in BCD (binary coded decimal) format, which is FORTRAN readable; sample printouts were also obtained to check results. The program MATCH is listed in Appendix B.

It should be noted that only the highway network was considered in developing relative accessibility variables. Although the trip generation models were for person-trips, using the highway network only would not introduce any appreciable bias in the case of IRTADS; mainly because transit passenger trips constituted only 4.1 percent of all the person-trips [70, p. 25]. Moreover, the transit in IRTADS area was entirely bus service on the city streets.

The Program "ACCESS." A computer program "ACCESS" was written for either the IBM 7094 or the CDC 6500. The program accepts friction factor tapes and the activity size variables as inputs and generates relative accessibility measures for each zone in the study area. A listing of ACCESS is presented in Appendix B. The sequence of computer programs and the data flow for calculating relative accessibilities are shown schematically in Figure 5.

With ten activity measures (the available independent variables) and six sets of friction factors (one for each trip purpose), sixty measures of relative accessibility could be generated. However, not all sixty possible combinations were generated; only those that were meaningful to the trip generation analysis were used. For example, the relative accessibility of a zone to retail floor area could be meaningful in conjunction with home-based shopping person-trip productions. Also, the relative accessibility of a zone to single family dwellings would be meaningful in

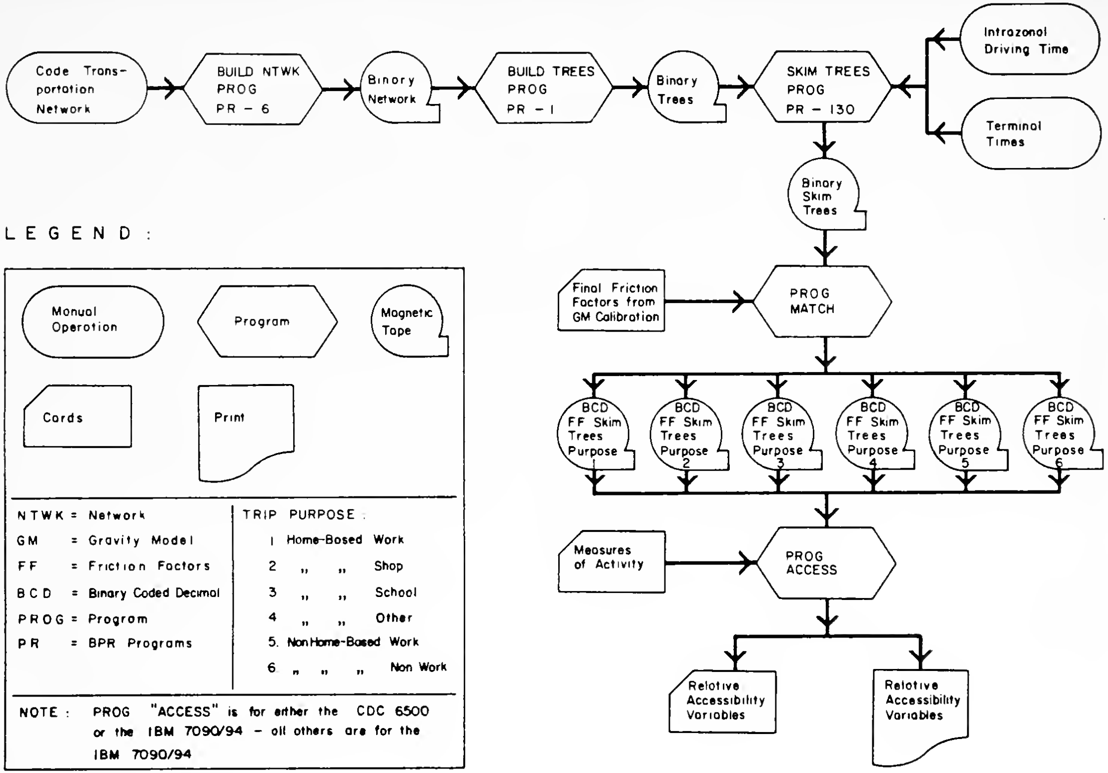


FIGURE 5 - THE PROCESS OF COMPUTING RELATIVE ACCESSIBILITIES

conjunction with home-based shop person-trip attractions. Twenty relative accessibility variables were generated and considered in the analysis; these are presented in Figure 6. It should be noted that the same measure of relative accessibility could be meaningful in conjunction with both the productions and attractions of some of the trip purposes.

A notation defining an accessibility variable should indicate the activity under consideration and the trip purpose whose set of friction factors were used in calculating that specific accessibility. A relative accessibility variable was abbreviated by a leading letter A denoting accessibility, followed by the abbreviation of the activity under consideration, and ended by a one digit number. The number, one to six, indicated the trip purpose of the friction factor used: home-based work, home-based shop, home-based school, home-based other, non home-based work-oriented, or non home-based non work-oriented respectively.

Delimiting the Central Area

As indicated in CHAPTER III, this investigation considered stratifying the study area into central and non central areas. It was assumed that the central and non central areas might reflect two different trip generation patterns due to the shape of the study area, its historical quasi-annular urban growth, and the configuration of the transportation system. This differentiation of the central and non central areas was categorical instead of numerical, and could best be treated through stratification.

MEASURES OF ACTIVITY	TRIP PURPOSE					
	HBWK 1	HBSHP 2	HBSCL 3	HBOTR 4	NHBWK 5	NHBNWK 6
Total Employment	P			P – A	P – A*	
Retail Employment						
Service Employment						
Retail Floor Area		P		P – A	P – A	P – A
Educational Floor Area			P	P – A	P – A	P – A
Dwelling Units		A		P – A		
Labor Force	A			P – A		
Population		A		P – A		
Cars				P – A		
Single Family Dwellings		A		P – A		

* P (and/or A) in cells indicates that the corresponding relative accessibility variable was considered in developing models of trip productions (and/or attractions) for the indicated trip purpose

FIGURE 6 – THE GENERATED RELATIVE ACCESSIBILITY VARIABLES

Areal Definitions of the Study Area. The IRTADS Study Area is bounded by a cordon encompassing all of Marion County and portions of Johnson and Hamilton Counties. The area within the cordon is divided into nine major subareas termed sectors (sectors zero through eight). Sector zero is the Central Business District (CBD); "...defined as the area within the inner loop." [69, p. 21]. The remaining eight sectors radiate outward from the CBD, dividing the Study Area into eight pie-shaped wedges.

The nine sectors are further divided into 72 districts. The districts are divided into 395 zones and the zones are further divided into 1200 subzones. Most of the analysis by IRTADS and by this investigation were carried on at the zone level. The delimitation of the central area, however, was carried at the district level; because the suggested procedure was not thought to be refined and precise enough to warrant an analysis on the zonal level. The general layout of the Study Area, and its sectors and districts are shown in Figure 7.

Criteria for Delimiting the Central Area. The rationale behind the procedure developed to delimit the central area was tied to the expected character and attributes of a central area. A larger proportion of the land in a central area is expected to be in urban use. A relatively small proportion of the land in a central area is expected to be devoted for residential uses. Conversely, a relatively high

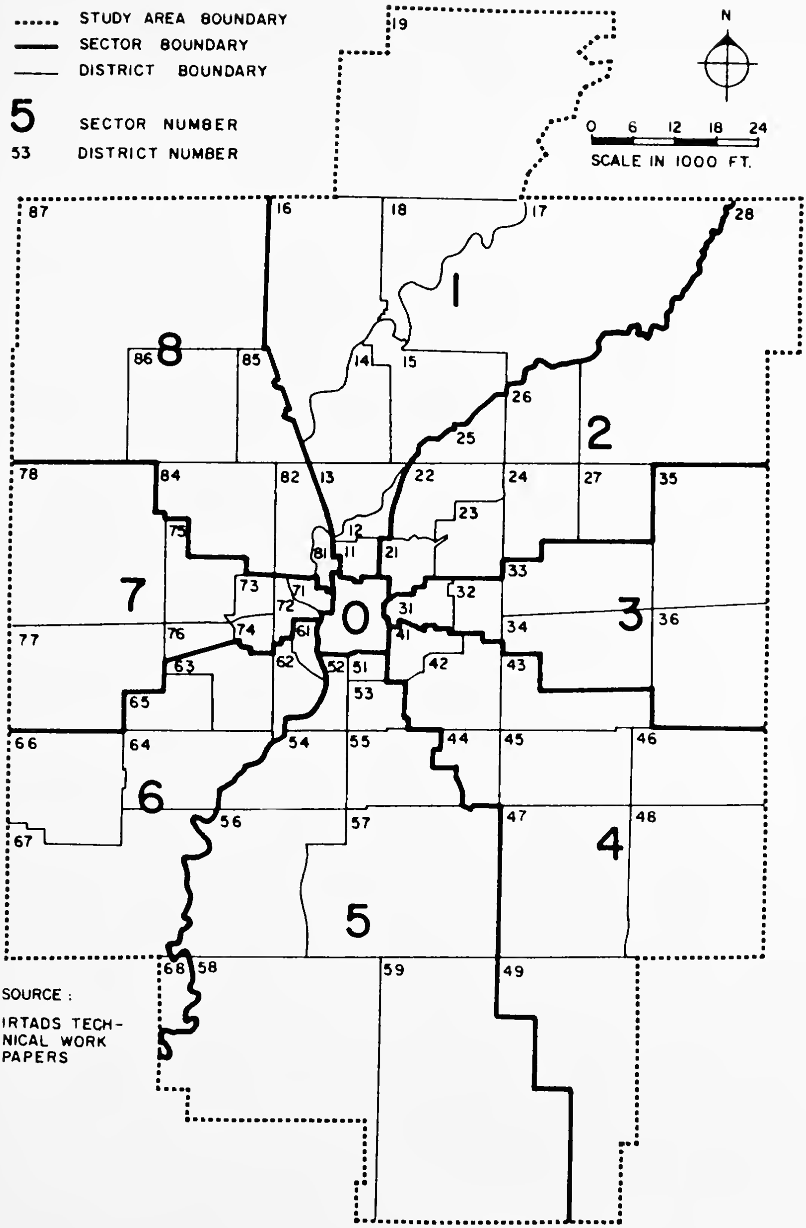


FIGURE 7 – STUDY AREA SECTORS AND DISTRICTS

proportion of the land in the central area was expected to be in uses which are known to seek central location.

Land was considered to be in urban use if it did not belong to any of the following classifications:

1. quarry and mining;
2. automobile junkyard;
3. water;
4. in agricultural use; or
5. vacant.

The percentage of urban land in residential use was the measure of the intensity of residential activity. The percentage of land in urban use was the measure of urbanity. Among the different trade and service uses on which IRTADS had floor area information, the following were chosen as uses that seek central location:

1. wholesale trade (without warehousing);
2. general retail trade;
3. auto retail;
4. apparel, furniture, and appliance retail;
5. retail use not otherwise classified;
6. finance, business, and professional services;
7. contract construction services;
8. governmental services;
9. personal services; and
10. services not otherwise classified [71, pp. 42-50].

Educational services were excluded because schools do not necessarily seek central locations. The floor area in

hundreds of square feet consumed by the above 10 uses in each district, per acre of land in urban use in the same district was calculated. This ratio measures not only the amount of these uses which are represented, but also reflects the intensity of use. If this ratio was unitless, that is, the numerator and denominator were of the same units, it would be indicative of the average number of floors per structure in each district.

Using the above three measures, the following conditions were set for delimiting the central area:

1. The delimitation was to be performed at the district level.
2. The central area would, probably, include all of sector zero and some of the qualifying surrounding districts.
3. The districts of the central area should all be contiguous and connected.
4. A district which would qualify must satisfy at least two of the three following criteria:
    a. In the lower quartile of all the districts of the study area in percentage of urban land in residential use.
    b. In the upper quartile of all the districts of the study area in the percentage of land in urban use.

c. In the upper quartile of all districts of the study area in the ratio of hundreds of square feet of uses usually seeking central location to acres of urban land in each district.

For each district, the values of the variables defining the three criteria were calculated. The values corresponding to criterion "a" were sorted in ascending order; the values for the other two criteria were sorted descendingly. Since the total number of districts is 72, then the first 18 districts corresponding to the sorted values would represent the lower quartile under criterion "a," and the upper quartiles under criterion "b" and criterion "c." To avoid any illusive precision, the values of each criterion corresponding to the 18th ordered district were rounded to the nearest integer. Thus, numerically, the cut-off points for the three criteria were determined to be:

Criterion a: percent of urban land in residential use less than, or equal to 32%.

Criterion b: percent of land in urban use greater than, or equal to 88%.

Criterion c: hundreds of square feet of selected trade and services per acre of urban land greater than, or equal to 8.

The supporting computations are presented in Table C1.

All the nine districts of sector zero and districts: 11, 21, 31, 41, 51, 52, 61, 71, and 72 satisfied the four conditions. District 81 satisfied only

criterion "a" of condition 4. However, it was decided to include it, in order to keep the central area completely connected.

A pictorial representation of the delimitation process is presented in Figure C1 of Appendix C. The study area as stratified into central and non central areas is shown in Figure 8. The districts of the central area constituted 105 zones out of the 395 in the study area.

Model Building

Guidelines for Model Building

Multiple linear regression models of trip generation were developed using the computer program "BMD-2R, stepwise regression" [72]. In addition to the desired statistical qualities of the developed models, other important factors were, also, considered.

Conceptual Validity. The consideration of relative accessibility was mainly to achieve a sounder conceptual basis for trip generation. In addition, only independent variables that were logically related to the specific dependent variable under consideration were allowed to enter when developing regression equations for that dependent variable. The causal-logical relationship was considered prior to the mere statistical correlation analysis. Association and correlation do not prove causality; causality should only be hypothesized on theoretical or conceptual grounds.

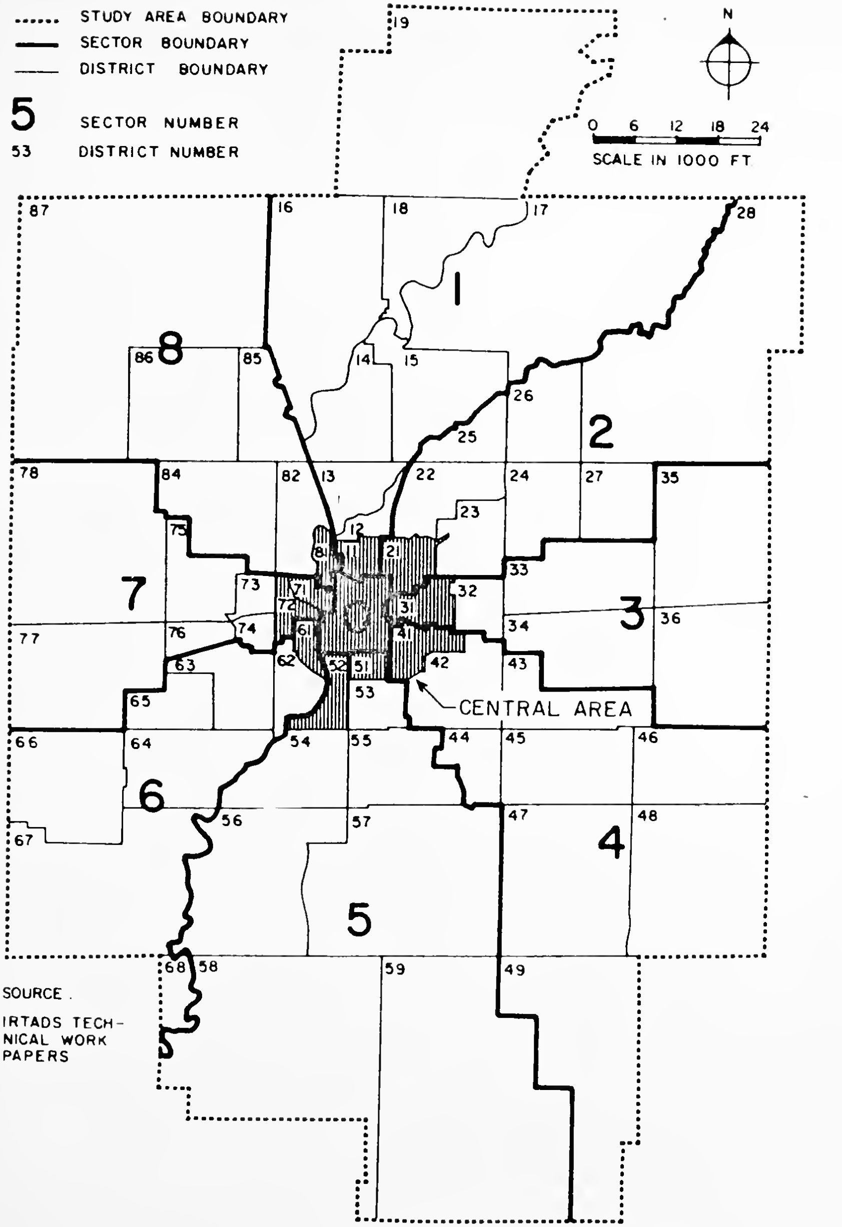


FIGURE 8 – STUDY AREA STRATIFICATION : CENTRAL AND NON CENTRAL

Model stability is one of the desirable products of conceptual validity. Relationships that are not conceptually valid, if established from today's data, are more apt not to hold in the future. Predictive equations of trip generation should hold for the future, in order to have any forecasting capability.

Another facet of conceptual validity is the sign of the regression coefficient. Because of collinearity in the variables, the coefficient of one of the independent variables could be contrary to the theoretical relationship and this condition might be statistically acceptable. In spite of this, it was decided to delete those variables whose coefficient had a sign contrary to conceptual expectations. This should increase the statistical validity of the model as it tends to reduce the effects of collinearity.

Simplicity. To keep the models as simple as possible, unnecessary transformations of, and interactions among the original independent variables were avoided. Interactions beyond the product of two independent variables were considered difficult to interpret; thus were avoided except if the third variable of the product was the dummy-variable defining the location of a zone in the central or non central areas.

Keeping the structure of the model as simple as possible by not going to higher order interactions curtails the propagation of measurement errors [73]. Another aspect of

simplicity is parameter parsimony. While it is valuable to include all the relevant independent variables, thus reducing specification errors, it is doubtful if it would be advantageous to do so when, as is the case for transportation studies, the input data are inherently plagued by measurement errors. As an emphasis of this research, the number of independent variables in the model was kept to a minimum.

Stability. In order that the developed models be stable over a time period, the prerequisite for allowing a variable to enter the model was a hypothesized causal relationship rather than a mere correlation.

Stability was also sought over the range of the values of the independent variables. This could be quite a difficult criterion to account for during model building. A study of the range of the independent variables for the forecast year was undertaken, and possible problem zones were identified. Recommendations will be made to ameliorate this condition.

Sensitivity. It is desirable that the response or the dependent variable be sensitive to changes in each of the independent variables in the model. The cost of adding one more independent variable would not be justifiable if the dependent variable is not sensitive to changes in the added independent variable. The sensitivity of the dependent variable to each of the independent variables in the model was

tested by calculating the standardized regression coefficients (the regression coefficients multiplied by the ratio of the standard deviation of the independent variable under consideration to the standard deviation of the dependent variable).

Statistical Considerations

Stepwise Regression. The computer program used by this research, as mentioned earlier, was "BMD-2R, stepwise regression" [72]. Several procedures are available to develop multiple regression models. The "tear-down" or "backward elimination" method starts with a model containing all the available independent variables, and subsequently eliminates some of the independent variables until a model with prescribed statistical features is reached. The "build-up" or "forward selection" procedure strives for a similar final outcome; however, working in the opposite direction by inserting one more independent variable at a time. Stepwise regression is an improved version of forward selection procedure [14]. The independent variables in the model are re-examined at the end of each step. The variable, that might have been the best single variable to enter at an earlier step, might prove to be unnecessary at a later stage because of the relationship between it and other variables now in the equation. Thus, at each step, the partial F-test

for each variable in the equation was evaluated and compared to a preselected percentage point of the appropriate F distribution. Stepwise regression evaluates the contribution of each independent variable in the model at the end of each step, regardless if the independent variable has entered at the last step or at any earlier step.

Partial F or Sequential F-Test. By far, the most important statistic in conjunction with regression analysis is the multiple coefficient of determination (R^2). It measures the proportion of total variability in the dependent variable explained by the regression model; R^2 varies between zero and one; a value of zero indicates a complete lack of fit, while a value of one implies a perfect correlation. In stepwise regression, a test is needed at each step to check if the increase in R^2 contributed by each added independent variable in the equation is significantly different than zero. The following F-statistic tests that the contribution of the k independent variables is significantly greater than zero.

$$F_{k,n-k-1} = \frac{R_k^2/k}{(1 - R_k^2)/(n - k - 1)}$$

where

n = number of observations,

k = number of independent variables, and

R_k^2 = coefficient of multiple determination of a model with k independent variables.

The calculated F-statistic is compared to a tabulated $F_{k,n-k-1,1-\alpha}$, where α is the probability of type I error, or the level of significance. The level of significance chosen should depend on the consequences of rejecting a true hypothesis. The level of significance for including a variable was set at 0.010, and for deleting a variable at 0.005. The selection of these values is based on acceptance of a relatively high risk of including a variable which does not belong. Once this variable has been accepted, there is a lower risk acceptable for its retention in the equation based on the entry of other independent variables.

The blind use of the F-test may result in developing a regression model which involves more independent variables than are of practical significance. In transportation studies, the number of observations is large resulting in an F-statistic which is statistically significant even when the absolute increase is R^2 is very small. The criterion of a significant increase in R^2 proved to be superfluous in the majority of the cases; as other criteria such as simplicity, parsimony, and reasonableness controlled the number of variables to be included in the model.

Standard Error of Estimate. Another statistic of interest is the standard error of the estimate (s). It is the square root of the residual mean square. The smaller the value of this statistic the more precise the predictions would be. The criterion of reducing s must be used

cautiously; since s can be made small by including enough parameters in the model, just as R^2 can be increased . As more independent variables are included in the equation, the decrease in s will be at a decreasing rate. Reduction of s is desirable if many degrees of freedom for error are remaining [14].

Another way of looking at the reduction in s is to consider it in relation to the dependent variable -- namely, as a percentage of the mean value of the dependent variable. Standard error of estimate as a percentage of the mean of the dependent variable is referred to as the coefficient of variation (C.V.).

t-Test on Regression Coefficients. It is sometimes desirable to test if each of the estimated regression parameters are significantly different than zero. The ratio of each regression coefficient to its standard error is distributed as student-t. If the regression coefficient of one of the independent variables does not pass the t-test, it can be deleted from the equation.

The three criteria of R^2, s, and significance of the regression coefficient are not independent. Usually, the decision can be made on the basis of R^2 alone.

Model Identification

The first set of models, that was developed by this investigation, was a rerun for each of the 13 dependent variables using the same independent variables established

by IRTADS for their equations. Some of IRTADS equations were developed using 389 zones out of the 395 [38]. There, probably, was a good reason for excluding some of the zones from the analysis by IRTADS. The excluded zones and the reason for their exclusion were not available for this investigation. Therefore, and in the interest of compatibility and comparability, data from all the 395 zones were used to re-estimate the parameters of the models developed by IRTADS. Those models, essentially developed by IRTADS, were used as a basis to compare with other developed models.

A second set of models that include relative accessibility variables were attempted for each of the 13 dependent variables. It was not, however, possible to develop satisfactory models for each of the 13 dependent variables. The first two sets of models were developed with data from the 395 zones with no distinction relative to location in the central or non central areas. Two more sets of models were developed; one corresponding to the set developed by IRTADS, the other to the set of models developed by this investigation. The models of these two sets contained a dummy-variable defining the location of a zone in the central or non central areas.

Thus, basically, four sets of models were developed. Two were without any relative accessibility variables among their independent variables; one of those two was the set developed by the traditional procedures for IRTADS; the

second set contained a dummy-variable which defined the zone location and/or some of the interaction of the dummy-variable with the other independent variables in the equation. Of the remaining two sets, each had relative accessibility variables and, in addition, one was calibrated with stratified data.

Figure 9 suggests a system to identify the developed models. Models for each dependent variable in every set were not possible to develop. Cells of Figure 9, where an abbreviation appears, indicate that an adequate model was developed for that dependent variable in the corresponding set.

Results of Model Development

The results of developing the four sets of trip generation models are presented below. The four sets will be referred to, as suggested in Figure 9, as follows:

1. Without accessibility, unstratified: Set W-U.
2. Without accessibility, stratified: Set W-S.
3. With accessibility, unstratified: Set A-U.
4. With accessibility, stratified: Set A-S.

Reference is made to Appendix A for the definition of the symbols that appear in the following discussion.

IRTADS Trip Generation Models: Set W-U

The results of re-estimating the parameters of the IRTADS models are presented only in summary form. The details of model development are not reported because it did not

THE DEPENDENT VARIABLES	WITHOUT ACCESSIBILITY W		WITH ACCESSIBILITY A	
	UNSTRATI_FIED U	STRATI_FIED S	UNSTRATI_FIED U	STRATI_FIED S
1. HOME_BASED WORK PERSON_TRIP PRODUCTIONS (HBWKP)	W_U_1	W_S_1	A_U_1	
2. ,, ,, SHOP ,, ,, ,, (HBSHPP)	W_U_2			
3. ,, ,, SCHOOL ,, ,, ,, (HBSCLP)	W_U_3	W_S_3	A_U_3	A_S_3
4. ,, ,, OTHER ,, ,, ,, (HBOTRP)	W_U_4			
5. NON HOME_BASED WORK_ORIENTED PERSON_TRIP PRODUCTIONS (NHBWKP)	W_U_5	W_S_5	A_U_5	A_S_5
6. NON HOME_BASED NON WORK_ORIENTED PERSON_TRIP ,, (NHBNWP)	W_U_6	W_S_6	A_U_6	A_S_6
7. TOTAL PERSON_TRIP PRODUCTIONS (TOTP)	W_U_7	W_S_7	A_U_7	A_S_7
8 HOME_BASED WORK PERSON_TRIP ATTRACTIONS (HBWKA)	W_U_8	W_S_8	A_U_B	A_S_B
9 ,, ,, SHOP ,, ,, ,, (HBSHPA)	W_U_9	W_S_9	A_U_9	A_S_9
10. ,, ,, OTHER ,, ,, ,, (HBOTRA)	W_U_10	W_S_10	A_U_10	A_S_10
11. NON HOME_BASED WORK_ORIENTED PERSON_TRIP ATTRACTIONS (NHBWKA)	W_U_11	W_S_11	A_U_11	A_S_11
12 NON HOME_BASED NON WORK_ORIENTED PERSON_TRIP ,, (NHBNWA)	W_U_12	W_S_12	A_U_12	A_S_12
13. TOTAL PERSON_TRIP ATTRACTIONS (TOTA)	W_U_13	W_S_13	A_U_13	A_S_13

INDICATES THAT NO SATISFACTORY MODEL WAS DEVELOPED.

FIGURE 9 - A SYSTEM TO IDENTIFY THE DEVELOPED TRIP GENERATION MODELS

involve decisions by this investigation. The model calibration was done for the same set of independent variables using the complete data from the 395 zones of the study area. Table 2 summarizes the set W-U, and Table 3 gives the summary statistics.

Activity-Accessibility Models: Set A-U

In developing trip generation models with accessibility variables, the independent variables which proved significant in developing set W-U were allowed to enter in addition to the appropriate relative accessibility variables; meaningful cross products of socio-economic variables and accessibility variables were also considered.

It is restated here that the number at the end of the abbreviation of an accessibility variable represented the trip purpose whose friction factor was used in generating the specific accessibility variable. The numbers one to six represented the following trip purposes: home-based work, home-based shop, home-based school, home-based other, non home-based work-oriented, or non home-based non work-oriented respectively.

Model A-U-1 (HBWKP). This model is summarized in Table 4. The only accessibility variable considered was the accessibility to total employment (AEMPTOT1). The cross products of AEMPTOT1 with labor force (LF) and with cars (CARS) were also considered. Only the cross product of

Table 2. IRTADS Trip Generation Models: Set W-U.

Dependent Variable	Constant	Coefficients of Independent Variables									
		Employment			Floor Area*						
		TOT	RTL	SRV	RTL	ED	DU	LF	POP	CARS	SFD
1. HBWKP	-14.423	--	--	--	--	--	--	1.524	--	--	--
2. HBSHPP	-29.431	--	--	--	--	--	--	--	--	0.708	0.306
3. HBSCLP	- 5.493	--	--	--	--	--	--	--	--	0.631	--
4. HBOTRP	58.519	--	--	--	--	--	--	--	0.033	1.540	--
5. NHBWKP	10.062	0.217	--	0.218	0.135	--	0.259	--	--	--	--
6. NHBNWP	23.472	--	0.548	0.178	0.257	0.105	--	--	--	0.346	--
7. TOTP	154.289	0.262	--	--	0.476	--	--	--	0.300	4.427	--
8. HBWKA	184.355	1.209	--	--	--	--	--	--	--	--	--
9. HBSHPA	146.982	--	1.691	--	0.604	--	--	--	--	--	--
10. HBOTRA	149.991	0.119	--	0.671	0.358	0.251	--	--	--	0.840	--
11. NHBWKA	15.192	0.182	--	0.194	0.182	--	0.261	--	--	--	--
12. HNBNWA	5.111	--	0.456	0.111	0.273	0.067	--	--	--	0.412	--
13. TOTA	199.137	1.530	2.930	--	1.317	1.124	--	--	--	2.153	--

* In hundreds of square feet.

Table 3. Summary Statistics of IRTADS Trip Generation Models: Set W-U.

	Trip Purpose	Model Identification	R-Squared	S.E.	Mean	C.V. (%)
1.	HBWKP	W-U-1	0.974	174.220	1069.002	16.30
2.	HBSHPP	W-U-2	0.887	227.484	592.633	38.38
3.	HBSCLP	W-U-3	0.630	344.894	429.608	80.28
4.	HBOTRP	W-U-4	0.903	379.002	1184.792	31.99
5.	NHBWKP	W-U-5	0.748	237.307	440.838	53.83
6.	NHBNWP	W-U-6	0.650	361.274	512.965	70.43
7.	TOTP	W-U-7	0.961	756.759	4229.838	17.89
8.	HBWKA	W-U-8	0.839	640.540	1166.081	54.93
9.	HBSHPA	W-U-9	0.442	976.162	615.316	158.64
10.	HBOTRA	W-U-10	0.679	635.271	1197.476	53.05
11.	NHBWKA	W-U-11	0.738	236.758	438.390	54.01
12.	NHBNWA	W-U-12	0.636	377.779	512.792	73.67
13.	TOTA	W-U-13	0.771	2037.703	4364.896	46.68

Table 4. Summary of Model A-U-1.

Dependent Variable -- Name: HBWKP
Mean: 1069.00
Standard Deviation: 1084.28
C.V. At Last Step (%): 15.52

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	-24.62451					
1	LF	1.31155	.04085	.9742	.9742	174.22	1
2	SFD	.27319	.04285	.9765	.0022	166.65	2
3	CARS* AEMPTOT1	.25625	.12150	.9768	.0003	165.92	3

AEMPTOT1 by CARS entered the model. Its contribution, as measured by the increase in R^2 and the reduction in the standard error of the estimate, is very little. Home-based work productions are expected to be inelastic to the condition of the transportation system. The trip from home to work is essential and regular. The frequency of this trip is least affected by the whims of the trip maker and the environmental conditions.

Model A-U-2 (HBSHPP). It was not possible to develop this model. The relative accessibility variable considered was that to retail floor area (ARTLFLR2). Neither ARTLFLR2 nor any of its cross products with population of a zone, dwelling units in a zone, single family dwellings in a zone, or the cars garaged in a zone proved significant. This investigation was not able to develop a model, of home-based shop trips from a zone, that had accessibility variables.

Model A-U-3 (HBSCLP). Table 5 summarizes this model. The only accessibility variable considered to be meaningful in developing a model of home-based school trip productions was the accessibility to educational floor area (AEDFLR3). Cross products of AEDFLR3 with school age population, dwelling units, single family dwellings, or cars were also considered. Only SFD*AEDFLR3 appeared in the chosen model, second to CARS, and with a negative coefficient. The negative sign of the coefficient could be understood if it is realized that a high accessibility to educational floor

Table 5. Summary of Model A-U-3.

Dependent Variable -- Name: HBSCLP
Mean: 429.61
Standard Deviation: 566.37
C.V. At Last Step (%): 75.95

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	- 7.79456					
1	CARS	0.78254	.03196	.6301	.6301	344.89	1
2	SFD* AEDFLR3	- 0.85128	.12407	.6698	.0397	326.30	2

indicates probably spatial proximity to schools; thus more school trips are apt to be walking trips.

Model A-U-4 (HBOTRP). It was not possible to develop this model. The accessibility variables considered in conjunction with home-based other trip productions were the accessibility to total employment, accessibility to retail floor area, and the accessibility to educational floor area. The cross products of each of the accessibility variables with the variables population, cars, or single family dwellings were also considered. None of the accessibility variables or its cross products could be satisfactorily included. Their contribution to the increase in R^2 or the reduction in the standard error of estimate would have been little. The signs of the regression coefficient of the accessibility variables would have been negative despite their positive first order correlation with the dependent variable because the relative accessibility variables were correlated to the independent variables population and cars.

Model A-U-5 (NHBWKP). This model is summarized in Table 6. Accessibilities to total employment, retail floor area, and educational floor area were considered. The cross products of the three accessibility variables by total employment, retail employment, and service employment respectively were also made available. Other cross products could have been considered, but because of the intercorrelation of most of the independent variables only those

Table 6. Summary of Model A-U-5.

Dependent Variable -- Name: NHBWKP
Mean: 440.84
Standard Deviation: 469.87
C.V. At Last Step (%): 53.89

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	34.78471					
1	EMPTOT* AEMPTOT5	.33086	.04955	.5012	.5012	332.26	1
2	RTLFLR	.15832	.01535	.6398	.1386	282.72	3
3	CARS	.19626	.01752	.7221	.0823	248.63	2
4	EMPTOT	.10307	.02070	.7365	.0144	242.43	4
5	EMPSRV	.15268	.03693	.7476	.0111	237.58	5

three were considered. The first variable to enter the equation was the cross product of total employment by accessibility to total employment. It also ranked highest as to the sensitivity of the dependent variable to the various independent variables in the model.

Model A-U-6 (NHBNWA). Table 7 presents a summary of the model. Only two accessibility variables were considered: the accessibility to retail floor area and the accessibility to educational floor area. Only those two were considered logical to investigate in conjunction with non home-based non work-oriented trip productions. The cross products of each of the accessibility variables by total employment, retail floor area, and educational floor area were also considered. In the developed model only the cross product of total employment by the accessibility to educational floor was present. It entered at the third step and ranked fourth as to the sensitivity of the dependent variable to the independent variables in the model.

Model A-U-7 (TOTP). This model is summarized in Table 8. All the independent variables used in the separate trip purpose production models were allowed to enter. Among the accessibility variables only the cross product of labor force by accessibility to total employment entered the final model in the fourth step. Its contribution to the statistical strength of the model is quite small.

Table 7. Summary of Model A-U-6.

Dependent Variable -- Name: NHBNWP
Mean: 512.96
Standard Deviation: 606.70
C.V. At Last Step (%): 71.10

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	-64.85785					
1	RTLFLR	.24929	.03858	.4205	.4205	462.42	2
2	CARS	.37479	.02712	.5916	.1710	388.72	1
3	EMPTOT* AEDFLR6	.20690	.05389	.6222	.0307	374.32	4
4	EMPRTL	.51210	.13392	.6357	.0135	368.04	3
5	AEDFLR6	386.97221	135.45393	.6432	.0075	364.71	5

Table 8. Summary of Model A-U-7.

Dependent Variable -- Name: TOTP
Mean: 4229.84
Standard Deviation: 3833.54
C.V. At Last Step (%): 18.61

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	202.6649					
1	CARS	5.0006	.0820	.9315	.9315	1004.88	1
2	RTLFLR	.5123	.0510	.9516	.0202	845.41	2
3	TOTEMP	.3618	.0353	.9577	.0061	791.04	3
4	LF* AEMPTOT1	.6591	.2972	.9583	.0005	787.11	4

Model A-U-8 (HBWKA). This model is presented in Table 9. Eleven independent variables were allowed to enter this model. The eleven variables were constituted of one accessibility variable: the accessibility to labor force, five socio-economic and demographic variables: total employment, retail employment, service employment, retail floor area, and educational floor area, and the five cross products of the accessibility variable with each of the socio-economic variables. The final model included two independent variables: total employment and the cross product of retail floor by the accessibility to labor force. The contribution of the latter is very little compared to the former.

Model A-U-9 (HBSHPA). The model is summarized in Table 10. The independent variables considered were three accessibility variables, three socio-economic and land use variables, and their nine cross products. The accessibilities to population, dwelling units, and single family dwellings were the three accessibility variables. Retail employment, service employment, and retail floor area were the socio-economic and land use variables. The chosen model had only two independent variables. The first variable to enter was the cross product of accessibility to single family dwellings by retail employment and the second variable was retail floor area. The variable including the accessibility measure entered first and contributed the most

Table 9. Summary of Model A-U-8.

Dependent Variable -- Name: HBWKA
Mean: 1166.08
Standard Deviation: 1592.38
C.V. At Last Step (%): 54.72

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	161.85608					
1	EMPTOT	1.19035	.02820	.8386	.8386	640.54	1
2	RTLFLR* ALF1	0.28603	.14227	.8403	.0016	638.07	2

Table 10. Summary of Model A-U-9

Dependent Variable -- Name: HBSHPA
Mean: 615.32
Standard Deviation: 1303.67
C.V. At Last Step (%): 141.99

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	89.93472					
1	ASFD2* EMPRTL	10.21393	.91122	.4955	.4955	927.11	1
2	RTLFLR	.48669	.06849	.5531	.0576	873.72	2

to the statistical strength of the model.

Model A-U-10 (HBOTRA). Table 11 presents a summary of this model. Four accessibility variables were considered: accessibility to population, accessibility to dwelling units, accessibility to single family dwellings, and accessibility to cars. The cross products of retail floor area and educational floor area by accessibilites to population, single family dwellings, and cars were also considered. The cross product of retail floor area by accessibility to single family dwellings entered at the sixth step and was the last independent variable to be included.

Model A-U-11 (NHBWKA). This model is summarized in Table 12. The accessibilities to total employment, retail floor area, and educational floor area together with their cross products by retail floor area and educational floor area were considered. The selected model of non home-based work-oriented trip attractions had seven independent variables. The cross product of total employment by accessibility to total employment entered in the fourth step, and the cross product of retail floor area by accessibility to total employment entered in the sixth step. Those two variables ranked third and sixth respectively in the sensitivity of the dependent variable.

Model A-U-12 (NHBNWA). This model is presented in Table 13. Among the independent variables allowed to enter

Table 11. Summary of Model A-U-10.

Dependent Variable -- Name: HBOTRA
Mean: 1197.48
Standard Deviation: 1114.19
C.V. At Last Step (%): 52.90

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	160.42837					
1	CARS	.82217	.04955	.3803	.3803	878.18	1
2	EMPSRV	.66756	.09306	.5348	.1539	762.30	2
3	RTLFLR	.23336	.08027	.6224	.0881	687.26	4
4	EDFLR	.24854	.03521	.6672	.0447	646.09	3
5	EMPTOT	.11882	.03132	.6790	.0119	635.27	5
6	RTLFLR* ASFD4	.49408	.02749	.6817	.0026	633.46	6

Table 12. Summary of Model A-U-11.

Dependent Variable -- Name: NHBWKA
Mean: 438.39
Standard Deviation: 460.09
C.V. At Last Step (%): 54.15

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	34.24864					
1	EMPTOT	.10595	.02394	.4186	.4186	351.27	2
2	RTLFLR	.11319	.03554	.6082	.1896	288.71	4
3	CARS	.20841	.01780	.6843	.0760	259.51	1
4	EMPTOT* AEMPTOT5	.20742	.06210	.7227	.0385	243.50	3
5	EMPSRY	.15651	.03730	.7331	.0103	239.23	5
6	RTLFLR* AEMPTOT5	.17927	.08024	.7364	.0033	238.04	6
7	EMPRTL	.15665	.03820	.7385	.0021	237.38	7

Table 13. Summary of Model A-U-12.

Dependent Variable -- Name: NHBNWA
Mean: 512.79
Standard Deviation: 622.47
C.V. At Last Step (%): 72.90

Step No.	Independent Variable	At Last Step Reported		R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
		Regression Coefficient	Standard Error of Coefficient				
	CONSTANT	-50.8158					
1	RTLFLR	.26187	.03956	.3894	.3894	487.02	3
2	CARS	.42478	.02789	.6101	.2207	389.66	2
3	EMPTOT* AEDFLR6	.13821	.05539	.6267	.0166	381.75	1
4	EMPRTL	.43686	.13732	.6358	.0091	377.57	4
5	EDFLR* AEDFLR6	.06832	.03660	.6430	.0072	374.30	5
6	AEDFLR6	217.93838	155.74181	.6448	.0018	373.84	6

were two accessibilities: to retail floor area, and to educational floor area; and their cross products by total employment, retail floor area, and educational floor area. The cross product of total employment by accessibility to educational floor area entered the model at the third step but ranked first as to the sensitivity of the dependent variable. The cross product of educational floor area by accessibility to educational floor area, and the variable of accessibility to educational floor area entered in the fifth and sixth steps respectively, and ranked fifth and sixth respectively.

Model A-U-13 (TOTA). Table 14 summarizes this model. All the accessibility variables used in developing models of trip attractions together with all the other independent variables were available. Many alternative models of total trip attractions could be built because of the inter-correlations in the independent variables. The selected model was a five-variable model. The variable entering at the fifth step was a cross product of retail employment by the accessibility to single family dwellings calculated using the friction factors for the shopping trips. The same variable ranked third as to the sensitivity of total trip attractions to the five independent variables in the model.

Table 14. Summary of Model A-U-13.

Dependent Variable -- Name: TOTA
Mean: 4364.90
Standard Deviation: 4235.03
C.V. At Last Step (%): 43.58

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	264.50121					
1	RTLFLR	1.17220	.15372	.3931	.3931	3303.40	5
2	EMPTOT	1.53025	.08600	.5355	.1424	2893.58	1
3	CARS	1.93847	.14641	.7047	.1692	2310.07	2
4	EDFLR	1.04277	.10462	.7625	.0578	2074.37	4
5	ASFD2* EMPRTL	17.40110	2.01197	.8008	.0383	1902.18	3

The Use of Dummy-Variables

Stratification of the study area into central and non central areas in the analysis required the use of a dummy-variable. The dummy-variable represents a qualitative variable which is unmeasurable on a continuous numerical scale [74]. This variable is given a value of one if the zone belongs to the central area and zero if the zone belongs to the non central area. Therefore, the coefficient of the dummy-variable represents a comparison of zones in the central area to zones in the non central area [75]. Stated differently, the coefficient of the dummy-variable measures the comparison in the impact of the zone location on the response variable.

A Geometrical Interpretation. The geometrical interpretation of the use of dummy-variables will be illustrated by an example. Consider the case of simple linear regression:

$$Y = a + b_1X_1 + b_2Z$$

where

Y = the dependent variable,

a = the constant term of the equation,

X_1 = a continuously measurable independent variable,

b_1 = the regression coefficient of X_1,

Z = a dummy-variable, it equals one for zones in the central area and zero for zones in the non central area, and

b_2 = the regression coefficient of the dummy-variable

Geometrically the preceding equation could be plotted in cartesian coordinates as illustrated in Figure 10.

Consider, now, the same regression equation with a cross product term. Thus,

$$Y = a + b_1X_1 + b_2Z + b_3X_1Z$$

where

Y, a, X_1, b_1, Z, and b_2		are the same as previously defined; and
b_3	=	the regression coefficient of the cross product term.

The above relationship could also be plotted in cartesian coordinates as shown in Figure 11.

As can be seen from Figures 10 and 11, a dummy-variable in the regression model connotes a shift in the relationship depending on the group to which the observation belongs. A dummy-variable and its cross products indicate both a shift in the intercept and a change in the slope. Of course, a regression model might contain the cross product term only together with the other independent variables, indicating a change in the slope without any shift in the intercept.

It should be noted that interactions of the dummy-variable with other independent variables present a problem; because leaving them out may reduce the model precision, but if included the model becomes unwieldy.

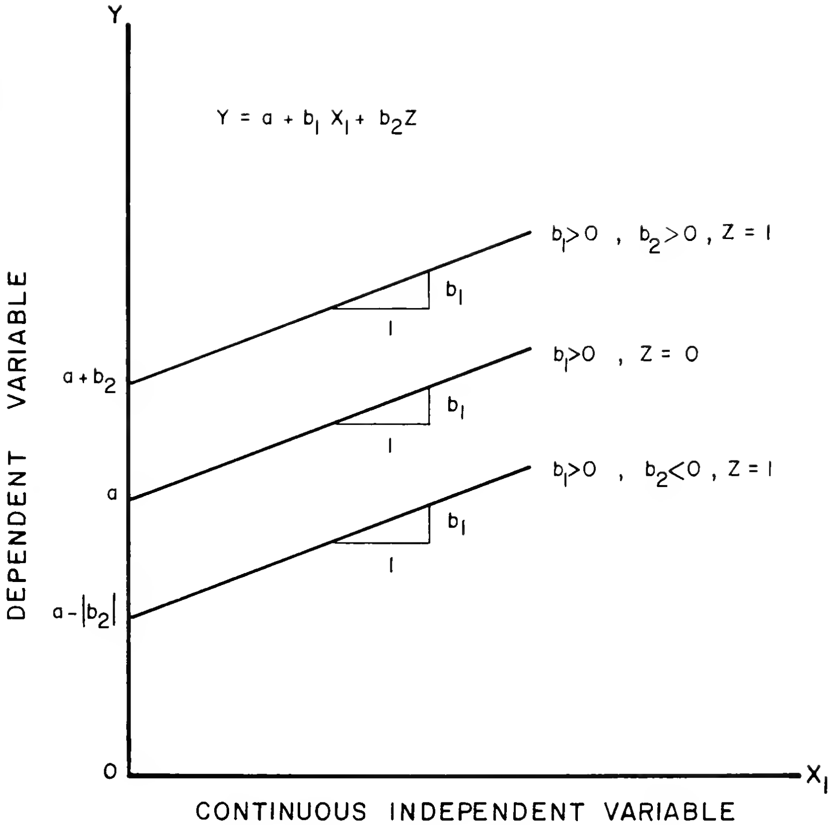


FIGURE 10 - LINEAR REGRESSION WITH A DUMMY VARIABLE

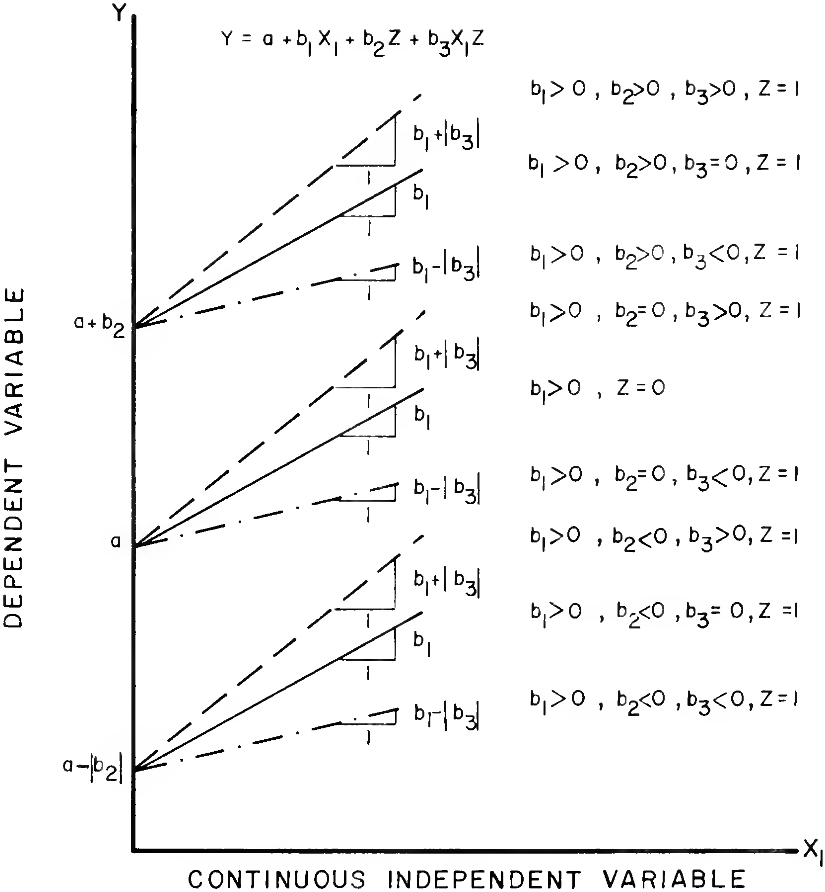


FIGURE 11 _ LINEAR REGRESSION WITH A DUMMY_ VARIABLE AND ITS CROSS PRODUCT

Statistical Tests of Significance. In order to establish a criterion for the statistical validity of stratifying the input data, statistical tests are needed. Stratified data were used for both the trip generation models of the IRTADS type and those developed by this investigation that included accessibility variables. The two sets with stratified data were referred to as: set W-S and set A-S respectively. The same independent variables that were chosen for sets W-U and A-U were allowed to enter plus the previously defined dummy-variable, and all the cross products of the dummy-variable by the other independent variables.

In order to establish a rule for limiting the size of a regression model, another more conservative criterion was utilized in addition to the partial-F test. The following statistical test on the significance of the net contribution of considering the factor of location was utilized. Consider the following F-ratio:

$$F = \frac{(R_A^2 - R_B^2)(n - k_1 - k_2 - 1)}{(1 - R_A^2)(k_1)}$$

where

k_1 = number of independent variables representing the factor of zone location,

k_2 = number of the independent variables other than those representing the factor of location,

n = number of observations,

R_A^2 = coefficient of multiple determination for equation with $(k_1 + k_2)$ variables, and

R_B^2 = coefficient of multiple determination for equation with k_2 variables [75].

"Again, the significance of the net relationship depends heavily on the number of observations and on the number of variables." [75, p. 381]. The level of significance was conservatively set at $\alpha = 0.0005$.

The t-test used for testing if an individual regression coefficient is significantly different from zero, is of low utility in the case of dummy-variables [75]. This should be clear because the choice of the values of one and zero to indicate if a zone belongs to the central or the non central areas respectively, is arbitrary. Testing whether the coefficient of a dummy-variable is significantly different from zero was in effect testing for the significance of the difference between the two classes.

IRTADS Models with Dummy-Variables: Set W-S

For each of the 13 models, the dummy-variables and each of its cross products by the independent variables were allowed to enter. Thus, for a model with k independent variables in set W-U, the corresponding model in set W-S could have at most 2k + 1 independent variables. In most of the models, however, the number of independent variables did not reach 2k + 1. The possibility of almost doubling the number of independent variables is not a detriment to

parsimony, because one component of each of the added cross products was already in the equation. The only added cost is due to the collection of the information necessary to classify zones into central and non central. Thus stratification is believed to enable a more efficient use of the other available independent variables.

Model W-S-1 (HBWKP). This model is summarized in Table 15. Only the cross product term was added. The negative sign of the cross product term could indicate that the rate of home-based work productions for a zone in the central area per labor force is lower than the corresponding rate for a non central zone.

Model W-S-2 (HBSHPP). This model was not developed because neither the dummy-variable nor any of its cross products increased R^2 significantly. A possible interpretation is that the rates of home-based shop productions per car or per single family dwelling were not significantly different in the central and non central areas.

Model W-S-3 (HBSCLP). The model which is summarized in Table 16 has only the cross product term added. The negative sign of the regression coefficient shows that home-based school productions per car in the non central area are, almost, twice that of the central area.

Model W-S-4 (HBOTRP). The model is summarized in Table 17. It does not correspond exactly to model W-U-4, which

Table 15. Summary of Model W-S-1.

Dependent Variable -- Name: HBWKP
Mean: 1069.00
Standard Deviation: 1084.28
C.V. At Last Step (%): 16.00

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	-10.60047					
1	LF	1.53323	.01258	.9742	.9742	174.22	1
2	LF*DUMMY	- .09296	.02602	.9751	.0008	171.67	2

Table 16. Summary of Model W-S-3.

Dependent Variable -- Name: HBSCLP
Mean: 429.61
Standard Deviation: 566.37
C.V. At Last Step (%): 77.14

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	17.99241					
1	CARS	.64092	.02349	.6301	.6301	344.89	1
2	CARS*DUMMY	- .39454	.06801	.6594	.0292	331.40	2

Table 17. Summary of Model W-S-4.

Dependent Variable -- Name: HBOTRP
Mean: 1184.79
Standard Deviation: 1214.34
C.V. At Last Step (%): 31.09

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	40.74936					
1	CARS	1.25985	.07933	.9026	.9026	379.42	1
2	SFD	.60087	.12628	.9077	.0051	369.77	2
3	SFD*DUMMY	.26418	.13245	.9087	.0009	368.37	3

used the variable population (POP). Instead, the variable single family dwelling (SFD) was used in W-S-4. The cross product of SFD and the dummy-variable was the only additional variable.

Model W-S-5 (NHBWKP). The model is presented in Table 18. The cross product of dwelling units by the dummy-variable and the dummy-variable itself were not included in the model. The other three cross products entered consecutively in the fifth, sixth and seventh steps. No interpretation of the signs of the regression coefficients will be attempted because the relationships are probably confounded due to the inter-correlation of the independent variables.

Model W-S-6 (NHBNWP). The model is summarized in Table 19. It includes cross products of the dummy-variable by retail floor area and by retail employment. Both of these had negative regression coefficients indicating a greater rate in the increase of non home-based non work-oriented productions due to an increase in retail floor area or retail employment in the central area than in the non central area. The dummy-variable itself entered at the last reported step with a positive coefficient, indicating a shift in the intercept in addition to the different slopes for zones in the central and non central areas. The positive sign of the regression coefficient of the dummy-variable could be interpreted that the location of a zone in the

Table 18. Summary of Model W-S-5.

Dependent Variable -- Name: NHBWKP
Mean: 440.84
Standard Deviation: 469.87
C.V. At Last Step (%): 51.89

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	11.72495					
1	EMPTOT	.18911	.01307	.5001	.5001	332.65	1
2	DU	.23485	.02137	.6690	.1689	271.01	4
3	RTLFLR	.18917	.02108	.7209	.0519	249.18	2
4	EMPSRY	.36660	.06533	.7467	.0258	237.70	3
5	RTLFLR* DUMMY	- .12791	.02928	.7528	.0061	235.10	7
6	EMPTOT* DUMMY	.10339	.02265	.7598	.0070	232.05	6
7	EMPSRY* DUMMY	- .27581	.07898	.7672	.0073	228.77	5

Table 19. Summary of Model W-S-6.

Dependent Variable -- Name: NHBNWP
Mean: 512.96
Standard Deviation: 606.70
C.V. At Last Step (%): 52.69

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	21.15990					
1	RTLFLR	.27697	.03632	.4193	.4193	462.90	3
2	RTLFLR* DUMMY	- .13378	.05817	.6040	.1847	382.76	6
3	CARS	.23699	.02258	.6725	.0685	348.55	4
4	EMPRTL	1.85357	.14878	.7090	.0366	328.93	1
5	EMPRTl * DUMMY	- 1.74294	.20548	.7484	.0394	306.26	2
6	EDFLR	.10160	.01496	.7812	.0328	285.96	5
7	EMPSRV	.21279	.03720	.8018	.0206	272.54	7
8	DUMMY	105.18589	38.30156	.8056	.0038	270.27	8

central area increases the potential of producing non home-based non work-oriented trips. However, this interpretation is very naive because of the complex relationship among the other seven independent variables. The dummy-variable in this model, actually, has a damping effect on the response of the model to changes in retail floor area and retail employment.

<u>Model W-S-7 (TOTP)</u>. This model is summarized in Table 20. The last two steps reported include the cross products of retail floor area and total employment by the dummy-variable. Again, any attempt to interpret the sign of the regression coefficients could be invalidated by the complex inter-correlations of the independent variables.

<u>Model W-S-8 (HBWKA)</u>. The model is summarized in Table 21. It shows that the rate of home-based work attractions per total employment is higher in the non central area than in the central area.

<u>Model W-S-9 (HBSHPA)</u>. Table 22 is a summary of this model. Two cross product terms are in the model. Both are with negative regression coefficients indicating the different drawing powers of activities in the central area compared to activities in the non central area.

<u>Model W-S-10 (HBOTRA)</u>. This model is summarized in Table 23. The only cross products included in the model were

Table 20. Summary of Model W-S-7.

Dependent Variable -- Name: TOTP
Mean: 4229.84
Standard Deviation: 3833.54
C.V. At Last Step (%): 17.15

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	180.59828					
1	CARS	4.34315	.12710	.9315	.9315	1004.88	1
2	RTLFLR	.75102	.06639	.9516	.0202	845.42	2
3	EMPTOT	.20630	.04033	.9518	.0061	790.97	5
4	POP	.29345	.04750	.9614	.0037	756.72	3
5	RTLFLR* DUMMY	- .56342	.09327	.9640	.0025	732.33	4
6	EMPTOT* DUMMY	.17639	.06021	.9647	.0008	725.29	6

Table 21. Summary of Model W-S-8.

Dependent Variable -- Name: HBWKA
Mean: 1166.08
Standard Deviation: 1592.38
C.V. At Last Step (%): 54.74

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	186.26332					
1	EMPTOT	1.24163	.03425	.8386	.8386	640.54	1
2	EMPTOT* DUMMY	- .06991	.04465	.8401	.0015	638.27	2

Table 22. Summary of Model W-S-9.

Dependent Variable -- Name: HBSHPA
Mean: 615.32
Standard Deviation: 1303.67
C.V. At Last Step (%): 115.63

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	46.35370					
1	RTLFLR	.73787	.09137	.4099	.4099	1002.74	2
2	RTLFLR* DUMMY	- .69793	.14229	.5909	.1811	835.93	4
3	EMPRTL	4.52315	.39101	.6817	.0908	738.31	1
4	EMPRTL* DUMMY	- 3.00903	.54002	.7052	.0235	711.48	3

Table 23. Summary of Model W-S-10.

Dependent Variable -- Name: HBOTRA
Mean: 1197.48
Standard Deviation: 1114.19
C.V. At Last Step (%): 50.04

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	191.96310					
1	CARS	.69924	.05004	.3804	.3804	878.18	1
2	EMPSRV	.78469	.09226	.5343	.1539	762.30	3
3	RTLFLR	.58755	.05380	.6224	.0881	687.26	2
4	EDFLR	.29882	.03902	.6672	.0447	646.09	4
5	RTLFLR* DUMMY	- .42897	.06948	.6979	.0307	616.33	5
6	EMPTOT	.12698	.02967	.7113	.0134	603.27	6
7	EDFLR* DUMMY	- .17796	.07096	.7159	.0046	599.20	7

those of the dummy-variable with each of retail floor area and educational floor area. The regression coefficients in both cases were negative. This indicates that attraction rates of different land uses are different, and that the rate for the same land use differs with location. In the presented model of home-based other attractions, the rates of trip attractions per unit of educational floor area or retail floor area were higher for non central locations than for central locations for constant values of other independent variables in the model.

Model W-S-11 (NHBWKA). The model is summarized in Table 24. The cross products of each of the four independent variables from set W-U by the dummy-variables appear in this model.

Model W-S-12 (NHBNWA). The model is summarized in Table 25. It included in addition to two cross product terms the dummy-variable itself. Model W-S-12 is similar to model W-S-6 in this respect. Only those two models of set W-S included the pure dummy-variable in addition to cross product terms indicating a shift in the intercept from central to non central locations in the non home-based non work-oriented trip purpose.

Model W-S-13 (TOTA). This model is summarized in Table 26. Three cross product terms appear in the model. It would be unsafe to interpret the signs of the regression

Table 24. Summary of Model W-S-11.

Dependent Variable -- Name: NHBWKA
Mean: 438.39
Standard Deviation: 460.09
C.V. At Last Step (%): 52.22

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	22.52875					
1	EMPTOT	.16459	.01309	.4192	.4192	351.07	2
2	DU	.20939	.02418	.6196	.2004	284.49	4
3	RTLFLR	.25065	.02191	.7158	.0962	246.23	1
4	EMPSRV	.35635	.06563	.7370	.0213	237.13	3
5	RTLFLR* DUMMY	- .14387	.03093	.7461	.0091	233.31	5
6	DU*DUMMY	.07996	.03937	.7494	.0033	232.10	8
7	EMPSRV* DUMMY	- .25516	.07915	.7528	.0034	230.87	6
8	EMPTOT* DUMMY	.06289	.02309	.7574	.0047	228.93	7

Table 25. Summary of Model W-S-12.

Dependent Variable -- Name: NHBNWA
Mean: 512.79
Standard Deviation: 622.47
C.V. At Last Step (%): 55.05

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	4.78930					
1	RTLFLR	.31396	.03794	.3888	.3888	487.24	3
2	RTLFLR* DUMMY	- .18938	.06076	.6157	.2269	386.86	5
3	CARS	.29518	.02358	.7071	.0914	338.14	4
4	EMPRTL	1.74866	.15540	.7375	.0304	320.53	1
5	EMPRTL* DUMMY	- 1.66464	.21462	.7713	.0338	299.57	2
6	EMPSRV	.14854	.03886	.7859	.0146	290.24	7
7	EDFLR	.06355	.01562	.7945	.0086	284.74	6
8	DUMMY	111.22555	40.00647	.7985	.0040	282.30	8

Table 26. Summary of Model W-S-13.

Dependent Variable -- Name: TOTA
Mean: 4364.90
Standard Deviation: 4235.03
C.V. At Last Step (%): 38.71

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	368.24188					
1	RTLFLR	1.60368	.22642	.3931	.3931	3303.40	4
2	EMPTOT	1.37475	.09615	.5356	.1425	2893.47	2
3	CARS	1.57595	.14069	.7046	.1690	2310.57	5
4	EDFLR	1.10643	.09312	.7622	.0576	2075.66	6
5	RTLFLR* DUMMY	- 1.08683	.36010	.8085	.0463	1865.08	7
6	EMPRTL	8.67254	.95076	.8297	.0212	1761.25	1
7	EMPRTL* DUMMY	- 7.12184	1.29638	.8411	.0114	1703.29	3
8	EMPTOT* DUMMY	.37995	.14164	.8440	.0029	1689.82	8

coefficients due to the inter-correlation of the independent variables.

Activity-Accessibility Models with Dummy-Variables: Set A-S

The models of set A-U that were developed with accessibility variables in them were recalibrated by allowing in addition to the independent variables already in the equation a dummy-variable defining location and its cross products by the other independent variables. Additional accessibility or other independent variables were only allowed as cross products involving the dummy-variables.

Model A-S-1 (HBWKP). It was not possible to develop this model. The model of home-based work productions was least possible to improve by introducing accessibility variables or by stratification.

Model A-S-2 (HBSHPP). This model was not developed because it was not possible to develop the corresponding model A-U-2.

Model A-S-3 (HBSCLP). The model as summarized in Table 27 indicates that the rate of home-based school productions per car in a zone is lower for the central area than for the non central area. The addition of the cross product of cars by the dummy-variable increased the R^2 of this model substantially.

Table 27. Summary of Model A-S-3.

Dependent Variable -- Name: HBSCLP
Mean: 429.61
Standard Deviation: 566.37
C.V. At Last Step (%): 72.66

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	15.49822					
1	CARS	.79164	.03061	.6301	.6301	344.39	1
2	SFD*AEDFLR3	- .84573	.11871	.6698	.0397	326.30	2
3	CARS* DUMMY	- .39105	.06407	.6985	.0287	312.18	3

Model A-S-4 (HBOTRP). This model was not developed because it was not possible to develop the corresponding model A-U-4.

Model A-S-5 (NHBWKP). This model is summarized in Table 28. The cross products of the dummy-variable by retail employment and service employment appear in the model and both with a negative regression coefficient. A three-term product of total employment by the accessibility to total employment by the dummy-variable appears in the model with a positive regression coefficient. After accounting for the other variables in the model, the rate of non home-based work productions per total employment by the accessibility to total employment in the central area is approximately three times that in the non central area.

Model A-S-6 (NHBNWP). A summary of the model appears in Table 29. The cross products of the dummy-variable by retail floor area, retail employment, service employment, and total employment are present in the model. It is almost futile to try to interpret the signs of the regression coefficients of the dummy-variable cross products because of the inter-correlation of the independent variables.

Model A-S-7 (TOTP). The model is summarized in Table 30. As can be seen, the contribution of accessibility variables and of the dummy-variable to the improvement of the model of total productions is quite small.

Table 28. Summary of Model A-S-5

Dependent Variable -- Name: NHBWKP
Mean: 440.84
Standard Deviation: 469.87
C.V. At Last Step (%): 49.41

Step No.	Independent Variable	At Last Step Reported		R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
		Regression Coefficient	Standard Error of Coefficient				
	CONSTANT	37.28197					
1	EMPTOT* AEMPTOT5	.16057	.05861	.5012	.5012	332.36	9
2	RTLFLR	.12727	.02331	.6398	.1386	282.72	8
3	CARS	.16396	.01727	.7221	.0823	248.63	7
4	EMPTOT	.11645	.02041	.7365	.0144	242.43	2
5	DUMMY* EMPRTL	- .72286	.10165	.7482	.0117	237.30	1
6	DUMMY*EMPTOT* AEMPTOT5	.31315	.05107	.7605	.0123	231.74	4
7	EMPRTL	.52034	.10847	.7743	.0139	227.23	5
8	EMPSRV	.35754	.06723	.7804	.0061	222.45	3
9	DUMMY* EMPSRV	- .33689	.08039	.7900	.0096	217.83	6

Table 29. Summary of Model A-S-6.

Dependent Variable -- Name: NHBNWP
Mean: 512.96
Standard Deviation: 606.70
C.V. At Last Step (%): 53.91

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	-47.76011					
1	RTLFLR	.29168	.03724	.4205	.4205	462.42	3
2	RTLFLR* DUMMY	- .17051	.05926	.6045	.1840	382.51	5
3	CARS	.26322	.02244	.6731	.0685	348.22	4
4	EMPTOT* AEDFLR6	.09296	.05408	.7143	.0412	325.95	9
5	EMPRTL	1.75406	.15352	.7502	.0359	305.19	1
6	EMPRTL* DUMMY	- 1.63408	.21200	.7819	.0317	285.53	2
7	AEDFLR6	418.10923	104.50898	.7878	.0059	281.97	6
8	EMPSRV* DUMMY	.13312	.05462	.7950	.0072	277.53	8
9	EMPTOT* DUMMY	.05230	.02683	.7970	.0020	276.53	7

Table 30. Summary of Model A-S-7.

Dependent Variable -- Name: TOTP
Mean: 4229.84
Standard Deviation: 3833.54
C.V. At Last Step (%): 17.83

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	219.84970					
1	CARS	4.89611	.08319	.9315	.9315	1004.88	1
2	RTLFLR	.78934	.06866	.9516	.0202	845.41	3
3	EMPTOT	.18763	.04192	.9577	.0061	791.04	4
4	DUMMY* RTLFLR	- .58331	.09698	.9600	.0023	770.31	2
5	DUMMY* EMPTOT	.21790	.06202	.9613	.0012	759.37	5
6	LF*AEMPTOT1	.72659	.28512	.9619	.0006	754.06	6

<u>Model A-S-8 (HBWKA)</u>. This model is summarized in Table 31. The cross product of total employment by the dummy-variable was added. It indicates that the rate of home-based work attractions per total employment is higher for zones of the non central area than for those of the central area. This is valid only after accounting for the third term in the model which is a cross product of retail floor area by the accessibility to labor force.

<u>Model A-S-9 (HBSHPA)</u>. As summarized in Table 32, the model includes the triple product of the dummy-variable by retail employment by the accessibility to single family dwellings. From the sign and value of the regression coefficients, it could be said that after accounting for retail floor area in a zone the attractive power of the zone to home-based shopping trips is increased if the zone lies in the non central area and is decreased if the zone lies in the central area as a function of the product of the accessibility of the zone to single family dwellings by the retail employment. To understand this relationship the results of model W-S-9 are recalled. The latter model showed that the attraction of shopping trips per unit of retail floor area or per retail employment was higher in the non central area than in the central area. Since the first order correlation of retail employment to retail floor area was relatively high (r = 0.810), this same relationship was, probably, reflected also in model A-S-9 in the sign and

Table 31. Summary of Model A-S-8.

Dependent Variable -- Name: HBWKA
Mean: 1166.08
Standard Deviation: 1592.38
C.V. At Last Step (%): 54.54

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	156.83679					
1	EMPTOT	1.22863	.03451	.8386	.8386	640.54	1
2	RTLFLR* ALF1	.32374	.14315	.8403	.0016	638.08	3
3	EMPTOT* DUMMY	- .08520	.04458	.8417	.0015	635.93	2

Table 32. Summary of Model A-S-9.

Dependent Variable -- Name: HBSHPA
Mean: 615.32
Standard Deviation: 1303.67
C.V. At Last Step (%): 127.92

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	92.07724					
1	ASFD2* EMPRTL	8.87862	.91826	.4955	.4955	927.11	4
2	RTLFLR	.52982	.08136	.5531	.0576	873.72	1
3	ASFD2*EMPRTL *DUMMY	-16.75028	1.74007	.6244	.0713	802.02	2
4	EMPRTL	1.35825	.34031	.6392	.0147	787.13	3

value of the regression coefficient of the variable that included the product of retail employment by the dummy-variable namely the triple product of relative accessibility to single family dwellings by retail employment by the dummy-variable.

<u>Model A-S-10 (HBOTRA)</u>. As summarized in Table 33, this model included three cross products of the dummy-variable. The sign and magnitude of the regression coefficients suggest that service employment and educational floor area are weak contributors to the attraction of home-based other trips in the zones of the central area.

<u>Model A-S-11 (NHBWKA)</u>. As shown in Table 34, the model contained cross products of the dummy-variable with three out of the ten independent variables in the model. In spite of the complex relationships among the independent variables in the model, the sign and magnitude of the regression coefficients suggest that retail employment and retail floor area have very little effect on the attractive power of a zone in the central area for non home-based work trips.

<u>Model A-S-12 (NHBNWA)</u>. This model is presented in Table 35. The same observation could be made about the effect of retail employment and retail floor area as noted above on model A-S-11. It is further observed that the interaction of total employment by the accessibility to

Table 33. Summary of Model A-S-10.

Dependent Variable -- Name: HBOTRA
Mean: 1197.48
Standard Deviation: 1114.19
C.V. At Last Step (%): 49.94

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	183.81310					
1	CARS	.69154	.05019	.3803	.3803	878.18	3
2	EMPSRV	.80146	.09278	.5314	.1510	764.67	4
3	RTLFLR	.78375	.11955	.6209	.0895	688.62	1
4	EDFLR	.30749	.03909	.6659	.0450	647.25	5
5	EMPSRV* DUMMY	- .80739	.16592	.6961	.0301	618.16	2
6	EMPTOT	.12259	.02963	.7093	.0132	605.35	8
7	EDFLR* DUMMY	- .21314	.07294	.7138	.0045	601.45	9
8	RTLFLR*ASFD4 *DUMMY	1.53731	.61425	.7162	.0024	599.66	6
9	RTLFLR*ASFD4	- .60177	.34105	.7185	.0023	598.03	7

Table 34. Summary of Model A-S-11.

Dependent Variable -- Name: NHBWKA
Mean: 438.39
Standard Deviation: 460.09
C.V. At Last Step (%): 49.84

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	44.79775					
1	EMPTOT	.10136	.02255	.4186	.4186	315.27	3
2	RTLFLR	.08798	.03445	.6082	.1896	288.71	10
3	CARS	.16935	.01744	.6843	.0760	259.51	4
4	EMPTOT* AEMPTOT5	.15092	.06244	.7227	.0385	243.50	6
5	RTLFLR* DUMMY	- .06840	.05131	.7361	.0134	237.85	9
6	EMPSRV	.14313	.03600	.7467	.0105	233.36	8
7	RTLFLR* AEMPTOT5	.28466	.08418	.7588	.0122	227.98	5
8	EMPRTL	.69173	.12382	.7671	.0082	224.34	1
9	EMPRTL* DUMMY	- .67476	.17377	.7753	.0083	220.60	2
10	EMPTOT*DUMMY* AEMPTOT5	.13136	.04500	.7802	.0049	218.48	7

Table 35. Summary of Model A-S-12.

Dependent Variable -- Name: NHBNWA
Mean: 512.79
Standard Deviation: 622.47
C.V. At Last Step (%): 54.70

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	- 40.35564					
1	RTLFLR	.32804	.03758	.3894	.3894	487.02	2
2	RTLFLR* DUMMY	- .22455	.05952	.6159	.2265	386.76	5
3	CARS	.30294	.02260	.7074	.0915	338.00	4
4	EMPRTL	1.69027	.15493	.7377	.0303	320.42	3
5	EMPRTL* DUMMY	- 1.57065	.21432	.7716	.0339	299.41	1
6	DUMMY*EMPTOT *AEDFLR6	.26267	.04975	.7940	.0224	284.72	6
7	AEDFLR6	351.67163	98.46892	.8005	.0066	280.50	7

educational floor area will increase the attractive power of zones in the central area for non home-based non work-oriented trips.

Model A-S-13 (TOTA). The model is summarized in Table 36. It would be difficult to draw conclusions on the effect of different independent variables in this model. However, it can be observed that the stratification of the data into central and non central contributes, probably, more to the soundness of the relationship than the inclusion of the accessibility variables.

Table 36. Summary of Model A-S-13.

Dependent Variable -- Name: TOTA
Mean: 4364.90
Standard Deviation: 4235.03
C.V. At Last Step (%): 38.27

Step No.	Independent Variable	At Last Step Reported: Regression Coefficient	At Last Step Reported: Standard Error of Coefficient	R^2 At Step Specified	Increase in R^2	Standard Error of Estimate	Sensitivity Rank
	CONSTANT	397.28546					
1	RTLFLR	1.58654	.22408	.3931	.3931	3303.40	4
2	EMPTOT	1.42140	.08328	.5355	.1424	2893.58	2
3	CARS	1.52154	.13779	.7047	.1692	2310.07	5
4	EDFLR	1.07661	.09272	.7625	.0578	2074.36	6
5	RTLFLR* DUMMY	- 1.12425	.35824	.8088	.0463	1863.66	7
6	EMPRTL	8.52290	.93179	.8300	.0212	1759.78	1
7	EMPRTL* DUMMY	- 7.95268	1.31890	.8414	.0115	1701.59	3
8	EMPRTL*ASFD2* DUMMY	14.37279	5.09711	.8451	.0037	1684.00	8
9	ALF1* EMPSRV	1.88680	.69961	.8480	.0029	1670.48	9

CHAPTER V. DISCUSSION OF RESULTS

Comparison of the Developed Models

Four sets of trip generation models were developed by this investigation. One of which was a recalibration of the set of models developed by IRTADs. The improvements introduced by each of the other three sets over the set developed by IRTADS, along the traditional approach, were investigated. Because the desirability of a model is determined by other than statistical criteria alone, the scope of the comparison was not restricted to statistical considerations. These criteria were discussed in Chapter IV.

Since similar criteria governed the building of the four sets of models, the statistical comparisons were limited to the two criteria of coefficients of multiple determination (R^2) and the coefficient of variation (C.V.). It was not always possible to conduct statistical tests of the significance of the increase in R^2 of the models of each set compared to the models of all others. Comparisons of two models were made where the independent variables of one were a subset of the independent variables of the other.

The significance of the increase in R^2 achieved in considering the factor of location of a zone could be tested

using an F-statistic as described in Chapter IV, however, this was not always possible. When the independent variabiles of one model are not a subset of the independent variables of the other model or when the two models do not have the same number of independent variables, the comparison, based upon changes in R^2, must be made entirely on personal judgment.

The scope of the comparisons included:

1. improvements achieved by introducing relative accessibility variables to the basic IRTADS models, that is, models of set A-U versus models of set W-U;
2. improvements achieved by calibrating the models with data stratified according to the zone location over the basic IRTADS models, that is, models of set W-S versus models of set W-U;
3. improvements achieved by calibrating the models containing relative accessibility variables with data stratified by location over similar models calibrated with unstratified data, that is, models of set A-S versus models of set A-U; and
4. improvements achieved by introducing both relative accessibility variables and calibrating the model with stratified data over the basic IRTADs models, that is, models of set A-S verus models of set W-U.

Comparison of Models with Accessibility Variables Versus Basic IRTADS Models (Set A-U Versus Set W-U)

Introducing relative accessibility variables to models of trip generation was proposed by this investigation to increase the conceptual validity of the models. It was hypothesized that trips generated by a zone are a function of the relative accessibility of the zone. In addition to the conceptual validity considerations, the statistical strength of the relationship was tested. Whenever the independent variables in a model of set W-U were a subset of the independent variables of the corresponding model of set A-U, the statistical test of the significance of the increase in R^2 was applicable. Table 37 summarizes the statistical test of the significance of the increase in R^2. The statistical test was not applicable for models where the trip purpose is the only entry in Table 37.

<u>HBWKP: Models with and without Accessibility</u>. Model A-U-1 included two more independent variables than model W-U-1, only one of the added independent variables was a relative accessibility measure. Thus the increase in R^2 attributed to it was considered in the test. The increase in R^2 was not significant at $\alpha = 0.0005$.

<u>HBSCLP: Models with and without Accessibility</u>. The increase in R^2 was significant at $\alpha = 0.0005$.

Table 37. Comparison of Models with Accessibility Variables Versus Basic IRTADS Models (Set A-U vs. Set W-U).

Trip Purpose	R_A^2	R_B^2	$R_A^2 - R_B^2$	k_1	$k_1 + k_2$	$n-k_1-k_2-1$	$1 - R_A^2$	F
1. HBWKP	.9768	.9765	.0003	1	3	391	.0232	5.0560
2. (HBSHPP)								
3. HBSCLP	.6698	.6301	.0397	1	2	392	.3302	47.1302*
4. (HBOTRP)								
5. NHBWKP#								
6. NHBNWP#								
7. TOTP#								
8. HBWKA	.8403	.8386	.0016	1	2	392	.1597	1.9637
9. HBSHPA#								
10. HBOTRA	.6817	.6790	.0119	1	5	388	.3183	14.5058*
11. NHBWKA#								
12. NHBNWA#								
13. TOTA#								

* Significant at $\alpha = 0.0005$

( ) No satisfactory models were developed.

\# Model developed but no statistical testing possible.

<u>NHBWKP: Models with and without Accessibility</u>. The addition of a relative accessibility variable did not increase the R^2, even with more independent variables in the equation. However, it is important to note that the relative accessibility independent variable was the first to enter, since it had the highest simple correlation with the dependent variable among all the independent variables which were considered for this model.

<u>NHBNWP: Models with and without Accessibility</u>. The model with relative accessibility variables did not achieve a higher R^2.

<u>TOTP: Models with and without Accessibility</u>. No increase in R^2 was achieved by including relative accessibility variables. The model W-U-7 is probably a more desirable model.

<u>HBWKA: Models with and without Accessibility</u>. The increase in R^2 attributed to including a relative accessibility variable was not significant at $\alpha = 0.0005$.

<u>HBSHPA: Models with and without Accessibility</u>. Including a relative accessibility variable increased the R^2 for the same number of independent variables, and the standard error of the estimate was reduced. A statistical test was not possible because the independent variables of one of the models were not a subset of the independent variables of the other.

<u>HBOTRA: Models with and without Accessibility</u>. The increase in R^2 due to introducing a relative accessibility variable was significant at $\alpha = 0.0005$.

<u>NHBWKA: Models with and without Accessibility</u>. Introducing relative accessibility variables did not increase R^2. The variable, dwelling units, in model W-U-11 was not allowed in model A-U-11 because a meaningful relationship with non home-based trips was not justifiable.

<u>NHBNWA: Models with and without Accessibility</u>. The increase in R^2 due to introducing relative accessibility variables was small. A statistical test on the significance of the increase was not possible, the independent variables of model W-U-12 were not a subset of the independent variables of model A-U-12.

<u>TOTA: Models with and without Accessibility</u>. The two models differ only in the last independent variable to enter. At the fifth step, model W-U-13 accepted the independent variable: retail employment (EMPRTL). At the same step, model A-U-13 accepted the variable: EMPRTL * ASFD2 which increased the R^2 by 0.0383 as compared to 0.0089 achieved by adding EMPRTL. The model A-U-13 is probably more desirable. The product of retail employment and the relative accessibility to single family dwellings (calculated using friction factors for the shopping trips) contributed more to the increase in R^2 than the variable retail employment alone.

Comparison of Stratified Models without Accessibility Variables Versus Basic IRTADS Models (Set W-S Versus Set W-U)

Table 38 summarizes the statistical test of significance described in CHAPTER IV. Stratification of the zones into central and non central areas increase the R^2 significantly at an α level of 0.0005 for all the models, except for the trip purpose of HBWKA.

Comparison of Stratified Models with Accessibility Variables Versus Unstratified Models with Accessibility Variables (Set A-S Versus Set A-U)

The purpose of the comparison was to assess the increase in R^2 achieved by stratifying the zones by location for trip generation models which included relative accessibility variables. The tests of significance are summarized in Table 39. Model A-S-1 was not possible to develop as were models A-U-2 and A-U-4. The increase in R^2 achieved by the models of set A-S over the models of set A-U were all significant at $\alpha = 0.0005$ except for HBWKA.

Comparison of Stratified Models with Accessibility Variables Versus Basic IRTADS Models (Set A-S Versus Set W-U)

The purpose of this comparison was to investigate the overall improvement attributed both to including relative accessibility variables and stratifying zones by location in set A-S over the basic IRTADS models of set W-U. Again, statistical tests were only possible if the independent variables of the model of set W-U were a subset of those of

Table 38. Comparisons of Stratified Models without Accessibility Variables Versus Basic IRTADS Models (Set W-S vs. Set W-U).

Trip Purpose	R_A^2	R_B^2	$R_A^2 - R_B^2$	k_1	$k_1 + k_2$	$n-k_1-k_2-1$	$1 - R_A^2$	F
1. HBWKP	.9751	.9742	.0009	1	2	392	.0249	14.1686*
2. (HBSHPP)								
3. HBSCLP	.6594	.6301	.0293	1	2	392	.3406	33.7217*
4. (HBOTRP)								
5. NHBWKP	.7672	.7467	.0205	3	7	387	.2328	11.3595*
6. NHBNWP	.8056	.6499	.1557	2	7	387	.1944	154.9791*
7. TOTP	.9647	.9614	.0033	2	6	388	.0353	18.1359*
8. HBWKA	.8401	.8386	.0015	1	2	392	.1599	3.6773
9. HBSHPA	.7052	.4422	.2630	2	4	390	.2948	173.9654*
10. HBOTRA	.7159	.6790	.0369	2	7	387	.2841	25.1325*
11. NHBWKA	.7574	.7370	.0204	4	8	386	.2426	8.1146*
12. NHBNWA	.7985	.6363	.1622	3	8	386	.2015	103.5718*
13. TOTA	.8440	.7714	.0726	3	8	386	.1560	59.8795*

* Significant at α = 0.0005

( ) No satisfactory models were developed.

Table 39. Comparison of Stratified Models with Accessibility Variables Versus Unstratified Models with Accessibility Variables (Set A-S vs. Set A-U).

Trip Purpose	R_A^2	R_B^2	$R_A^2 - R_B^2$	k_1	$k_1 + k_2$	$n-k_1-k_2-1$	$1 - R_A^2$	F
1. (HBWKP)								
2. (HBSHPP)								
3. HBSCLP	.6985	.6698	.0287	1	3	391	.3015	37.2196*
4. (HBOTRP)								
5. NHBWKP	.7900	.7476	.0424	4	9	385	.2100	19.4333*
6. NHBNWP	.7970	.6432	.1538	4	9	385	.2030	72.9224*
7. TOTP	.9619	.9583	.0036	2	5	389	.0381	18.3779*
8. HBWKA	.8417	.8403	.0014	1	3	391	.1583	3.4580
9. HBSHPA	.6392	.5534	.0858	1	3	391	.3608	92.9817*
10. HBOTRA	.7185	.6817	.0368	3	9	385	.2815	16.7768*
11. NHBWKA	.7802	.7385	.0417	3	10	384	.2198	23.9569*
12. NHBNWA	.8005	.6448	.1557	3	7	387	.1995	99.1227*
13. TOTA	.9613	.9583	.0030	2	5	389	.0387	15.0775*

* Significant at $\alpha = 0.0005$

( )No satisfactory models were developed.

the corresponding model of set A-S. Table 40 summarizes the statistical test of significance.

Models A-S-1, A-S-2, and A-S-4 were not possible to develop. Models where a statistical test was applicable are shown with complete entries in Table 40. The increase in R^2 was significant at the α level of 0.0005 for the following trip purposes: HBSCLP, HBSHPA, HBOTRA, and TOTA. No significant increase in R^2 was achieved for HBWKA. Those models for which a statistical test did not apply are discussed below.

Models of NHBWKP. Model A-S-5 included nine independent variables, five of which were socio-economic variables. Model W-U-5 included four socio-economic variables which were not a subset of the independent variables of model A-S-5. The increase in R^2 due to including four independent variables, measuring relative accessibility and specifying zone location, in addition to the five socio-economic variables in model A-S-5 was significant at $\alpha = 0.0005$.

Models of NHBNWP. Model A-S-6 had a higher R^2 and lower S.E. than those of model A-U-6. However, a statistical test of the significance of the increase of R^2 was not possible to perform.

Models of TOTP. The two models are identical in the first three steps. At the fourth step, model W-U-7 added the variable population while model A-S-7 added 3 more

Table 40. Comparison of Stratified Models with Accessibility Variables Versus Basic IRTADS Models (Set A-S vs. Set W-U).

Trip Purpose	R_A^2	R_B^2	$R_A^2 - R_B^2$	k_1	$k_1 + k_2$	$n-k_1-k_2-1$	$1 - R_A^2$	F
1. (HBWKP)								
2. (HBSHPP)								
3. HBSCLP	.6985	.6301	.0684	1	3	390	.3015	97.3254*
4. (HBOTRP)								
5. NHBWKP#								
6. NHBNWP#								
7. TOTP#								
8. HBWKA	.8417	.8386	.0031	2	3	389	.1583	3.8090
9. HBSHPA	.6392	.4422	.1970	2	4	388	.3608	105.9257*
10. HBOTRA	.7185	.6790	.0395	4	9	381	.2815	13.3654*
11. NHBWKA#								
12. NHBNWA#								
13. TOTA	.8480	.7714	.0766	4	9	381	.1520	48.0010*

* Significant at $\alpha = 0.0005$

( ) No satisfactory models were developed.

\# Models developed but no statistical testing possible.

variables at the fourth, fifth, and sixth steps. The combined effect of the increase in R^2 due to those 3 added variables was approximately equivalent to including the variable population. Thus model A-S-7 could not be considered as an improvement over model W-U-7.

Models of NHBWKA. In developing model A-S-11, the independent variable dwelling units, which entered second in model W-U-11, was excluded because the trip purpose being modeled was non home-based work, and neither ends of the trip would be expected to start or terminate at a dwelling unit. The contribution of the variable dwelling units to the R^2 of model W-U-11 was large. Even though model A-S-11 with 10 independent variables had a higher R^2 than model W-U-11 with 4 independent variables, no statistical test of significance could be used.

Models of NHBNWA. Model A-S-12 with 7 independent variables achieved a higher R^2 than model W-U-12 with 5. The increase could not be statistically tested.

Overall Comparisons

Values of R^2, C.V., and the number of independent variables for each model of the four developed sets are summarized in Table 41. A discussion of the improvements, if any, introduced to the models of the thirteen trip purposes follows.

Table 41. Comparative Summary Statistics: All Developed Models.

Model / Trip Purpose	R^2				C.V.(%)				Number of Independent Variables in Model			
	W-U	W-S	A-U	A-S	W-U	W-S	A-U	A-S	W-U	W-S	A-U	A-S
1 HBWKP	.974	.975	.977	*	16.30	16.00	15.52	*	1	2	3	*
2 HBSHPP	.887	*	*	*	38.38	*	*	*	2	*	*	*
3 HBSCLP	.630	.659	.670	.699	80.28	77.14	75.95	72.66	1	2	2	3
4 HBOTRP	.903	*	*	*	31.99	*	*	*	2	*	*	*
5 NHBWKP	.748	.767	.748	.790	53.83	51.89	53.89	49.41	4	7	5	9
6 NHBNWP	.650	.806	.643	.797	70.43	52.69	71.10	53.91	5	8	5	9
7 TOTP	.961	.965	.758	.962	17.89	17.15	18.61	17.83	4	6	4	6
8 HBWKA	.839	.840	.840	.842	54.93	54.74	54.72	54.54	1	2	2	3
9 HBSHPA	.442	.705	.553	.639	158.64	115.63	141.99	127.92	2	4	2	4
10 HBOTRA	.679	.716	.682	.719	53.05	50.04	52.90	49.94	5	7	6	9
11 NHBWKA	.738	.757	.739	.780	54.01	52.22	54.15	49.84	4	8	7	10
12 NHBNWA	.636	.796	.645	.801	73.67	55.05	72.90	54.70	5	8	6	7
13 TOTA	.771	.844	.801	.848	46.68	38.71	43.58	38.27	5	8	5	9

* No satisfactory model developed.

Models of HBWKP. Introducing relative accessibility variables, and/or stratifying the zones by location did not make any practical improvement over the model of set W-U with a single independent variable.

Models of HBSHPP. No acceptable alternatives to model W-U-2 were possible to develop.

Models of HBSCLP. The increase in R^2 was significant at α = 0.0005 for models W-S-3, A-U-3, and A-S-3 over W-U-3. Including relative accessibility variables, stratification by zone location, and a combination of those two actions improved the modeling of this trip purpose.

Models of HBCTRP. No acceptable alternatives to model W-U-4 were possible to develop.

Models of NHBWKP. There was a significant increase in R^2 attributed to stratification of zones by location at α = 0.0005 for both models W-S-5 and A-S-5 over W-U-5 and A-U-5 respectively. Adding only relative accessibility variables did not improve the model. The improvement due to stratification was more evident.

Models of NHBNWP. Stratification increased R^2 significantly in models W-S-6 and A-S-6 over models W-U-6 and A-U-6 respectively. Introducing relative accessibility variables did not contribute as much as stratification did.

<u>Models of TOTP</u>. Both stratification and relative accessibility variables did not drastically improve other models over W-U-7.

<u>Models of HBWKA</u>. Both stratification and relative accessibility variables did not drastically improve other models over W-U-8.

<u>Models of HBSHPA</u>. This trip purpose had the weakest model in the IRTADS set. Stratification strengthened the model but relative accessibility variables did not contribute as much.

<u>Models of HBOTRA</u>. It was possible to increase R^2 significantly by either stratifying with respect to location, or including relative accessibility variables, or both.

<u>Models of NHBWKA</u>. Stratification of both models W-U-11 and A-U-11 increased R^2 significantly. Including relative accessibility variables only, without stratification, did not improve the model.

<u>Models of NHBNWA</u>. The increase in R^2 of model W-S-12 over model W-U-12 was significant. Adding relative accessibility variables without stratification did not increase R^2 as much as stratification alone achieved. However, a seven independent variable model with both relative accessibility variables and stratification was slightly better than the eight independent variable model with stratification only.

Models of TOTA. Stratification increased R^2 significantly for both models with and without relative accessibility variables. Model A-S-13 with both relative accessibility variables and stratification increased R^2 significantly over model W-U-13.

CHAPTER VI. AN APPLICATION OF THE DEVELOPED MODELS

The four sets of developed models were solved with 1985 forecast of the independent variables. The productions and attractions forecasted for each zone by purpose, as predicted by each of the four sets of models, were compared to detect trends and establish if they were significantly different from each other. The forecasts of the socio-economic variables, and the minimum time skim trees of the proposed future network were obtained from IRTADS. The procedure described in CHAPTER IV was used to generate the 1985 relative accessibility variables that entered the regression equations.

A regression model applies in the region of values covered by the range of each of the independent variables in the model. In the case of multiple regression the inference space could be, indeed, quite restricted; especially, when considering the joint space defined simultaneously by the range of all the independent variables in the model. Defining this inference space could be an unwieldy combinatorial problem.

In the case of trip generation analysis with data aggregated at the zone level, the applicability of the regression models for forecasting might be reduced if some of

the zones expect growth that results in forecast of independent variables outside the range of the base year values. In addition to range, the distribution of the values of an independent variable in the base year and the forecast year could be of importance to the applicability of the regression model. The strength of a model lies in the region where most of the values of the independent variables fall. Theoretically, independent variables with uniform distributions provide regression models of equal strength over the whole range of the values of the independent variables.

1985 Forecast of the Socio-Economic Variables

The following observations about the forecast values of the socio-economic variables were by no means extensive or complete. The values of each of the ten socio-economic variables for each zone were tabulated in frequency distributions for the base year and the forecast year. It was observed that the shape of the distributions was similar for the two years; however, the distribution of the 1985 values was shifted to the right, that is, less frequent in the low range of values and more frequent in the high range of values. This is to be expected, it is because of the urban growth of the area. Appendix D presents the tabulated frequency distributions and the corresponding histograms.

The values of the 1985 forecasted socio-economic variables in each zone were checked against the ranges of 1964 values. The summary of this check is presented in Table 42.

Table 42. Values of 1985 Forecast of Socio-Economic Variables which are Outside the Range of the 1964 Values.

Zone Number	EMPTOT	EMPRTL	EMPSRV	RTLFLR	EDFLR	DU	_F	POP	CARS	SFD
72					9343					
82					6851					
100						4220	4820	15157	8090	3208
102						5413	5772	18149	10620	2947
108									4546	
110						4116	4260	13394	7223	3324
170	8463									
172									4897	
175						7322	7474	22159	9621	3769
177						7323	7773	23046	12390	4505
209					13522	17545	22212	57621	25265	13039
232									4387	
240									4271	3022
314						5268	6514	19311	7349	4257
324										2796
326					14108					
327					18894					
328					7873					
361					7330					
364									4597	
382							4050			
390							4387	12645	5314	
391							4802	13840	6285	
Highest Value, 1964	8239	3045	4173	6376	6838	3915	3993	12097	4222	2713

It shows the zones where values of the 1985 forecast socio-economic variables are outside the range of the 1964 values. Twenty three zones of the 395 zones of the study area had between one and five of the socio-economic variables that had forecasted values outside the range of 1964 values. The total employment forecasted value for 1985 exceeded the 1964 range for only zone 170. The educational floor area forecasted values for 1985 exceeded the 1964 range for each of zones 72, 82, 326, 327, 328, and 361.

It should be noted that only three zones out of 23 belong to the central area. Those were zones 326, 327, and 328, which are adjacent and constitute district number 71. The forecasted educational floor area in each of them was outside the range of 1964 values.

A third group of zones constituted of zones 100, 102, 108, 110, 175, 177, 209, 232, 240, 314, 324, 364, 382, 390, and 391 had forecasted values outside the range of the values of demographic variables that characterize residential areas.

The latter group of zones were observed to be at the borders of the study area, and were of very large land area compared to the other zones of the study area. The percentage of land in urban use for all the zones of this group was below 50% which indicates underdevelopment in 1964, but high potential for growth by 1985. The high growth potential and the large area of those zones resulted in a forecast of socio-economic variables outside the range of 1964 data.

Other 1985 Independent Variables

The developed models included seven relative accessibility variables. They were:

1. ALF1;
2. AEMPTOT1;
3. ASFD2;
4. AEDFLR3;
5. ASFD4;
6. AEMPTOT5; and
7. AEDFLR6.

Those seven variables were generated for 1985 using the procedure described in CHAPTER IV. The values of the friction factors were assumed to stay the same for the forecast year. Preliminary research has indicated that such an assumption was warranted [76].

The definition of the central and non central areas was assumed the same for the forecast year as for the base year. A check was not possible because forecast values for all the variables that were used to reach a stratification were not available. The procedure described in CHAPTER IV depended essentially on the ranking of the study area districts with respect to three measures of land use and land use intensity. The underlying assumption in using the same stratification for 1985 was that the ranking of the districts with respect to the three decision variables would be, essentially, the same in 1985. The exact ranking, of course,

was not the criterion; rather, the belonging to the upper or lower quartile. Thus the implied assumption was not over demanding.

1985 Productions and Attractions Forecast

Computer programs were written to solve the developed regression equations of trip generation for the forecast values of the independent variables. Four sets of forecasts for each of 13 trip purposes for every zone in the study area were obtained. In the event that the forecast value happened to be negative, due to a negative constant term in the equation, the forecast value was set to zero. The programs supplied outputs in the form of tabulated printouts and punched cards.

The results of the forecasts, by trip purpose for each of the zones of the study area, were compared against each other. The objective was to test if the four sets of models resulted in significantly different forecasts. Five forecast comparisons were considered:

1. models without accessibility and no stratification (basic IRTADS models) versus models with accessility and no stratification (set W-U versus set A-U);
2. models without accessibility and no stratification (basic IRTADS models) versus models without accessibility but with stratification (set W-U versus set W-S);

3. models with accessibility and no stratification versus models with accessibility and stratification (set A-U versus set A-S);
4. models without accessibility but with stratification versus models with accessibility and stratification (set W-S versus set (A-S); and
5. models without accessibility and no stratification (basic IRTADS models) versus models with accessibility and stratification (set W-U versus set A-S).

Each of the above comparisons were made for all the 395 zones of the study area as one group, for the 105 zones of the central area as a second group, and for the 290 zones of the non central area as a third group. In total, twelve comparisons were conducted.

The Paired t-Test

The statistical test used for the above comparisons was the paired t-test. In any one comparison two sets of forecasts, of equal size for each trip purpose, were under consideration. The two sets to be compared were denoted by Y and Y'. Each component of the two sets can be logically paired with respect to zone number.

The hypotheses to be tested were:

H_0: $\mu_Y = \mu_{Y'}$

H_1: $\mu_Y \neq \mu_{Y'}$

The set of the differences of the paired observations was defined as $D = Y - Y'$, or otherwise stated as $d_i = y_i - y_i'$ where $i = 1$ to n. The above hypotheses could be stated:

$$H_o: \mu_D = 0$$
$$H_1: \mu_D \neq 0$$

Let $\bar{d}$ and $s_{\bar{d}}$ be the estimate of μ_D and the standard deviation of the estimate of μ_D respectively, then

$$t = \frac{\bar{d}}{s_{\bar{d}}}$$

would be the test statistic. The above hypotheses were rejected if

$$|t| \geq t_{(1-\alpha/2)(n-1)}.$$

It could be shown that a paired t-test is equivalent to a randomized complete block design with two blocks. Equivalent conclusions could thus be reached using analysis of variance techniques.

A two-tailed test was used because there were no *a priori* reasons to suspect that the forecasts by the models of one of the sets would yield a higher or lower forecast than that by the models of another specific set.

Result of the Tests

The mean value of the differences between the pair of zonal forecast, for each trip purpose, by models of the two

compared sets and the associated t-statistic are tabulated in Appendix E. The test results are discussed below.

All Zones of the Study Area

A summary of the results is presented in Figure 12. The five comparisons are discussed below.

Forecasts by Basic IRTADS Models Versus Forecasts by Models with Accessibility Variables (Set W-U Versus Set A-U). Among the trip purposes compared three models did not forecast significantly different zonal productions and attractions. These three trip purposes were: HBSHPA, NHBNWA, and TOTA. It was observed that the forecast for all trip productions were significantly different at $\alpha = 0.05$, and on the average, models with accessibility variables consistently forecasted zonal productions larger than those by basic IRTADS models.

Forecasts by Basic IRTADS Models Versus Forecasts by Stratified Models without Accessibility Variables (Set W-U Versus Set W-S). For only two of the trip purposes, namely HBWKA and HBSHPA, the differences in the zonal forecasts, on the average, were significant at $\alpha = 0.05$. It is observed again that the forecasts by the stratified models were, on the average, larger than the forecasts by the basic IRTADS models, whenever the difference was significant.

TRIP PURPOSE	COMPARISON OF FORECASTS BY VARIOUS MODEL SETS									
	W-U vs. A-U		W-U vs. W-S		A-U vs. A-S		W-S vs. A-S		W-U vs. A-S	
	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$
1. HBWKP	SIG	−	NS	−						
2. HBSHPP										
3. HBSCLP	SIG	−	NS	+	NS	+	SIG	−	SIG	−
4. HBOTRP										
5. NHBWKP	SIG	−	NS	−	NS	+	SIG	−	SIG	+
6. NHBNWP	SIG	−	NS	+	NS	+	SIG	+	SIG	−
7. TOTP	SIG	−	NS	+	NS	+	SIG	−	SIG	−
8. HBWKA	SIG	−	SIG	−	NS	−	SIG	−	SIG	+
9. HBSHPA	NS	−	SIG	−	NS	+	SIG	+	SIG	+
10. HBOTRA	SIG	+	NS	+	SIG	−	SIG	−	SIG	−
11. NHBWKA	SIG	−	NS	−	NS	+	SIG	−	SIG	+
12. NHBNWA	NS	+	NS	+	SIG	+	SIG	+	SIG	+
13. TOTA	NS	+	NS	+	NS	−	NS	−	SIG	−

SIG $H_0 : \mu_Y = \mu_{Y'}$ was rejected at $\alpha = 0.05$
NS $H_0 : \mu_Y = \mu_{Y'}$ was not rejected at $\alpha = 0.05$
+ $\bar{d}$ for indicated purpose was positive
− $\bar{d}$ " " " " negative

indicates that either or both models do not exist

FIGURE 12 – PAIRED t_TEST : SUMMARY OF RESULTS ALL ZONES

Forecasts by Unstratified Models with Accessibility Variables Versus Forecasts by Stratified Models with Accessibility Variables (Set A-U Versus Set A-S). Forecasts of only two of the trip purposes were significantly different at $\alpha = 0.05$. The two trip purposes were HBOTRA and NHBNWA.

Forecast by Stratified Models without Accessibility Variables Versus Forecasts by Stratified Models with Accessibility Variables (Set W-S Versus Set A-S). All trip purposes except TOTA had significantly ($\alpha = 0.05$) different forecasts. No generalization was possible regarding the sign of the mean of the difference.

Forecasts by Basic IRTADS Models Versus Forecasts by Stratified Models with Accessibility Variables (Set W-U Versus Set A-S). The mean differences between the forecasts were significant at $\alpha = 0.05$ for all the trip purposes tested. No generalizations, however, were possible to make regarding the sign of the mean of the difference.

Zones of the Central Area

The same tests were conducted for the zones of the central area only. A summary of the test results is presented in Figure 13. The five comparisons are discussed below.

TRIP PURPOSE	COMPARISON OF FORECASTS BY VARIOUS MODEL SETS									
	W-U vs. A-U		W-U vs. W-S		A-U vs. A-S		W-S vs. A-S		W-U vs. A-S	
	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$
1. HBWKP	NS	−	SIG	+	X	X	X	X	X	X
2. HBSHPP	X	X	X	X	X	X	X	X	X	X
3. HBSCLP	NS	−	SIG	+	SIG	+	NS	−	SIG	+
4. HBOTRP	X	X	X	X	X	X	X	X	X	X
5. NHBWKP	NS	−	NS	−	NS	+	NS	−	SIG	+
6. NHBNWP	NS	−	SIG	+	SIG	+	NS	+	SIG	+
7. TOTP	SIG	−	NS	+	NS	+	SIG	−	SIG	−
8. HBWKA	NS	−	SIG	+	SIG	+	NS	−	SIG	+
9. HBSHPA	NS	−	SIG	+	SIG	+	NS	+	SIG	+
10. HBOTRA	NS	+	SIG	+	NS	−	SIG	−	SIG	−
11. NHBWKA	NS	−	NS	+	SIG	+	NS	+	SIG	+
12. NHBNWA	NS	−	SIG	+	SIG	+	NS	−	SIG	+
13. TOTA	NS	−	SIG	+	SIG	+	NS	−	SIG	−

SIG $H_0 : \mu_Y = \mu_{Y'}$ was rejected at $\alpha = 0.05$
NS $H_0 : \mu_Y = \mu_{Y'}$ was not rejected at $\alpha = 0.05$
\+ $\bar{d}$ for indicated purpose was positive
− $\bar{d}$ " " " " negative
X indicates that either or both models do not exist

FIGURE 13 — PAIRED t_ TEST : SUMMARY OF RESULTS

ZONES OF THE CENTRAL AREA

Forecasts by Basic IRTADS Models Versus Forecasts by Models with Accessibility Variables (Set W-U Versus Set A-U). Forecasts of total productions were significantly different at $\alpha = 0.05$, while differences between forecasts of all other trip purposes were not significant. Models with relative accessibility variables resulted in forecasts close, on the average, to those by IRTADS models in the zones of the central area.

Forecasts by Basic IRTADS Models Versus Forecasts by Stratified Models without Accessibility Variables (Set W-U Versus Set W-S). All the compared trip purposes had significantly different forecasts, except for the NHBWKP trip purpose. It was observed that for the trip purposes where the differences were significant, the means of the differences for zones of the central area were positive. Therefore, the basic IRTADS models would yield, on the average, forecasts in the central area higher than the forecasts by the corresponding stratified models.

Forecasts by Unstratified Models with Accessibility Variables Versus Forecasting by Stratified Models with Accessibility Variables (Set A-U Versus Set A-S). All the differences were significant except for the HBOTRA trip purpose. Again, the mean differences for all the trip purposes that had a significant difference were positive. Thus, stratification, probably, has the same effect on models with

relative accessibility variables as on models without them. Stratification tends to result in models that predict on the average lower trip productions and attractions for the zones of the central area.

Forecasts by Stratified Models without Accessibility Variables Versus Forecasts by Stratified Models with Accessibility Variables (Set W-S Versus Set A-S). Forecasts by models of the two compared sets were significantly different only for the two trip purposes: TOTP and HBOTRA. The sign of the mean difference for the latter two purposes was negative. Because only two trip purposes are involved, it is difficult to draw meaningful inferences.

Forecasts by Basic IRTADS Models Versus Forecasts by Stratified Models with Accessibility Variables (Set W-U Versus Set A-S). The zonal forecasts were significantly different on the average. It also appeared that forecasts by the traditional models of IRTADS would result in forecasts higher than those predicted by the models with relative accessibility and stratified data for all but one of the trip purposes, for zones of the central area. One would expect that future growth of the urban area would cause a decrease in the relative importance of the central area, and as such it would attract a lower proportion of the trips.

Zones of the Non Central Area

The comparison of the forecasts for the zones of the non central area are discussed below. The summary of the results of the difference tests are presented in Figure 14.

Forecasts by Basic IRTADS Models Versus Forecasts by Models with Accessibility Variables (Set W-U Versus Set A-U). The differences between the forecasts by the models of those two sets were significant at $\alpha = 0.05$ for all but one of the trip purposes, namely HBSHPA. No trends were detected in the sign of the mean of the differences.

Forecasts by Basic IRTADS Models Versus Forecasts by Stratified Models without Accessibility Variables (Set W-U Versus W-S). Forecasts of HBWKP, HBSCLP, NHBWKP, HBWKA, HBSHPA and HBOTRA by the models of the two sets were significantly different at $\alpha = 0.05$. It was also observed that the mean differences for all the above trip purposes were negative. This indicated a probable trend that models calibrated with stratified data predicted, on the average, zonal productions and attractions higher than the predictions by the basic IRTADS models, for the above trip purposes, in the zones of the non central area.

TRIP PURPOSE	COMPARISON OF FORECASTS BY VARIOUS MODEL SETS									
	W-U vs. A-U		W-U vs. W-S		A-U vs. A-S		W-S vs. A-S		W-U vs. A-S	
	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$	TEST	$\bar{d}$
1. HBWKP	SIG	–	SIG	–	X	X	X	X	X	X
2. HBSHPP	X	X	X	X	X	X	X	X	X	X
3. HBSCLP	SIG	–	SIG	–	SIG	–	SIG	–	SIG	–
4. HBOTRP	X	X	X	X	X	X	X	X	X	X
5. NHBWKP	SIG	–	SIG	–	NS	–	SIG	–	SIG	+
6. NHBNWP	SIG	+	NS	–	NS	–	SIG	+	SIG	–
7. TOTP	SIG	–	NS	–	NS	+	SIG	–	SIG	–
8. HBWKA	SIG	–	SIG	–	SIG	–	SIG	–	SIG	+
9. HBSHPA	NS	–	SIG	–	SIG	–	SIG	+	SIG	+
10. HBOTRA	SIG	+	SIG	–	SIG	–	NS	–	SIG	–
11. NHBWKA	SIG	–	NS	–	NS	+	SIG	–	SIG	+
12. NHBNWA	SIG	+	NS	–	NS	–	SIG	+	SIG	+
13. TOTA	SIG	+	NS	–	SIG	–	SIG	+	SIG	–

SIG $H_0: \mu_Y = \mu_{Y'}$ was rejected at $\alpha = 0.05$
NS $H_0: \mu_Y = \mu_{Y'}$ was not rejected at $\alpha = 0.05$
+ $\bar{d}$ for indicated purpose was positive
– $\bar{d}$ " " " " negative

X indicates that either or both models do not exist

FIGURE 14 – PAIRED t_TEST : SUMMARY OF RESULTS
ZONES OF THE NON CENTRAL AREA

Forecasts by Unstratified Models with Accessibility Variables Versus Forecasts by Stratified Models with Accessibility Variables (Set A-U Versus Set A-S). The models of the two sets resulted in forecasts that were significantly different at $\alpha = 0.05$ for five trip purposes, namely, HBSCLP, HBWKA, HBSHPA, HBOTRA, and TOTA. The mean of the difference was negative for the five above purposes. This substantiated the observation that models calibrated with stratified data tend to predict higher trips, on the average, for zones of the non central area.

Forecasts by Stratified Models without Accessibility Variables Versus Forecasts by Stratified Models with Accessibility Variables (Set W-S Versus Set A-S). Forecasts by the models of the two sets were significantly different at $\alpha = 0.05$ for all but the HBOTRA trip purpose.

Forecasts by Basic IRTADS Models Versus Forecasts by Stratified Models with Accessibility Variables (Set W-U Versus Set A-S). Forecasts for all the trip purposes as predicted by the models of the two sets were significantly different at $\alpha = 0.05$.

It can be noted from Figure 15 that, in the non central area, the basic IRTADS models (set W-U) overforecasted compared to stratified models with accessibility (set A-S), on the average, trips for the following five purposes: NHBWKP, HBWKA, HBSHPA, NHBWKA, and NHBNWA. For these trip purposes

only, those zones which did not individually have overforecasted trips are shown in Figure 15. These zones tend to fall in the vicinity and along the corridors defined by major thoroughfares. In other words the proposed accessibility stratified models (set A-S) tend to forecast greater numbers of trips in zones in the non central area for most trip purposes; and for the five purposes noted above, the models still forecast greater numbers where the zones have greater accessibility as provided by the corridors.

The Proposed Trip Generation Process

The trip generation models proposed by this research are functions of the status of the transportation system. In an operational transportation study future forecasts of trip generation would then be affected by the nature of the proposed transportation network. And since the proposed network should be designed to serve future trip generation; therefore, an iterative process should be followed. It would be terminated when an equilibrium between the future supply of transportation (proposed plan) and the demand for transportation (travel forecast) is reached. This iterative process is illustrated in Figure 16.

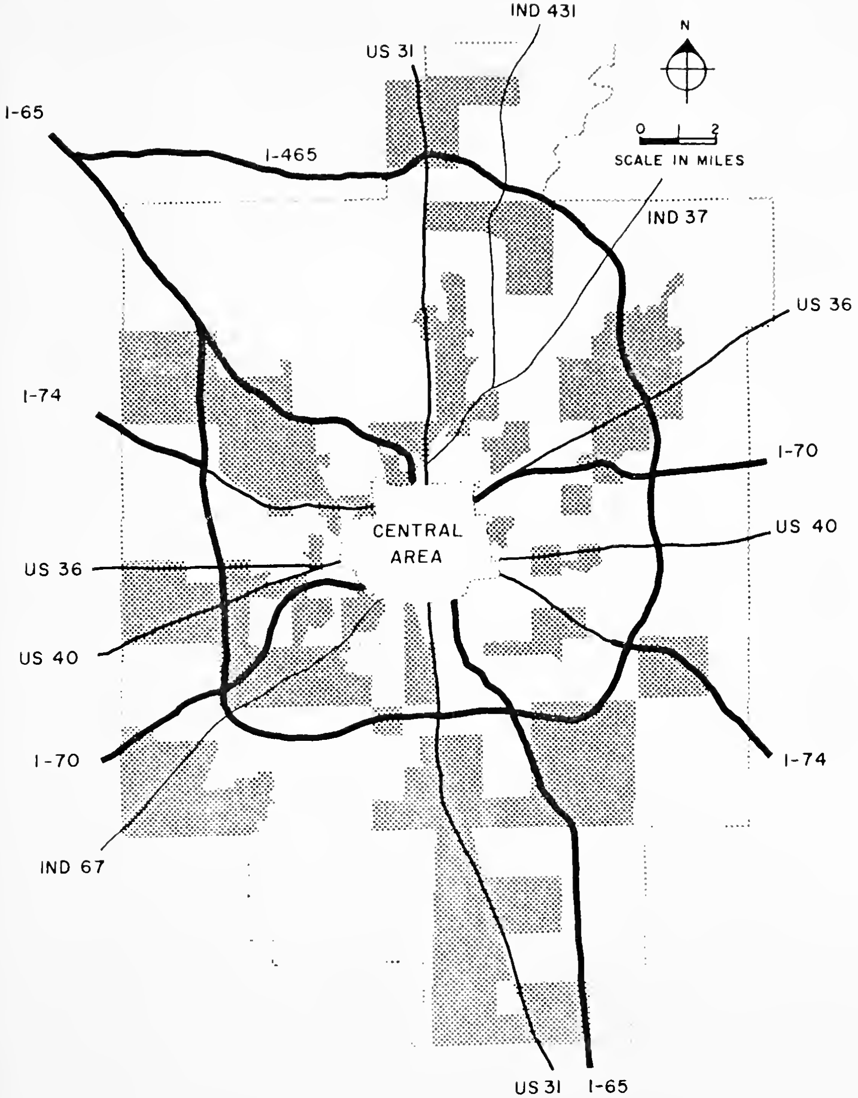


FIGURE 15 - NON CENTRAL AREA ZONES WHERE BASIC IRTADS MODELS DID NOT OVERFORECAST TRIPS FOR : NHBWKP, HBWKA, HBSHPA, AND / OR NHBWKA TRIP PURPOSES

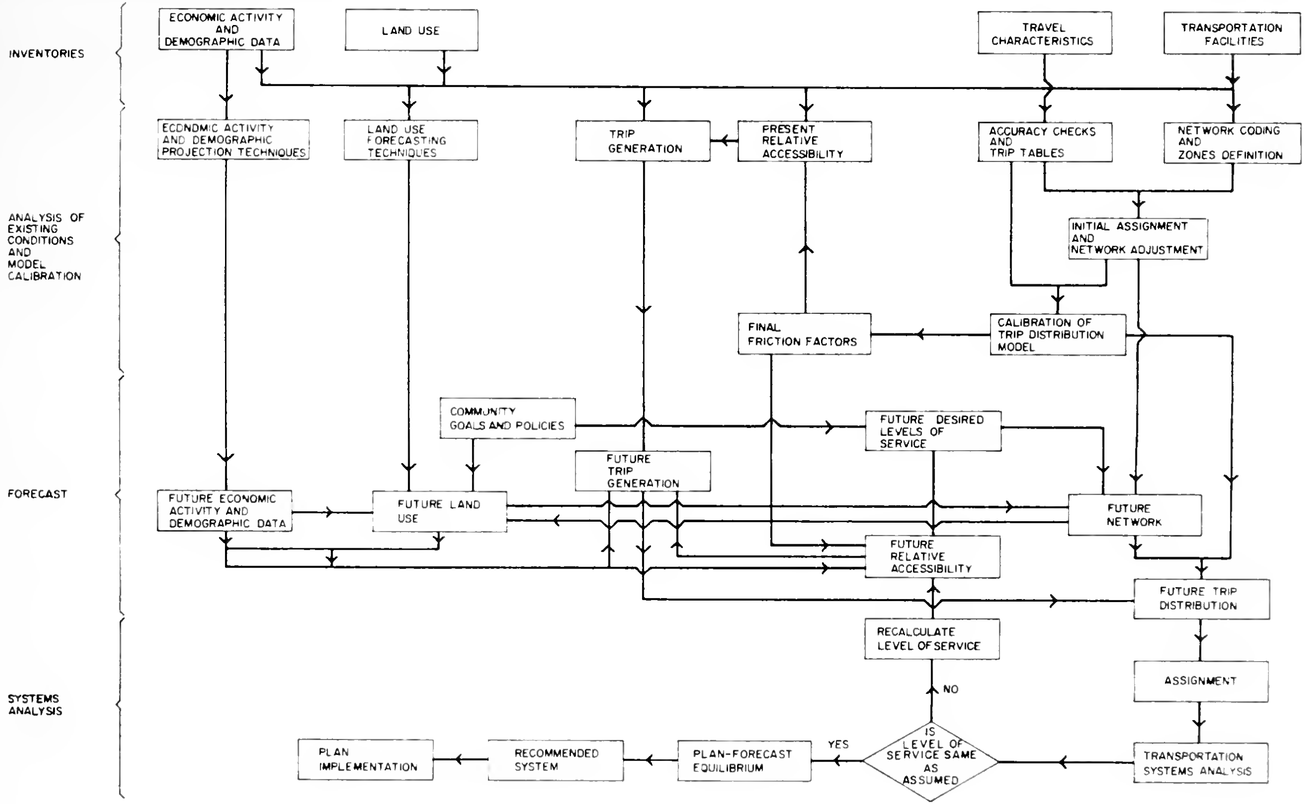


FIGURE 16 - THE PROPOSED TRIP GENERATION PROCESS

CHAPTER VII. CONCLUSIONS

Based on the preceding results and analyses the following conclusions can be drawn.

1. Among all of the relative accessibility variables considered, the following variables were included in the trip generation models which were developed:
   a. accessibility to employment in conjunction with home-based work person-trip productions and non home-based work-oriented person-trip attractions;
   b. accessibility to labor force in conjunction with home-based work person-trip attractions;
   c. accessibility to single family dwellings in conjunction with home-based shop person-trip attractions and home-based other person-trip attractions; and
   d. accessibility to educational floor area in conjunction with home-based school productions, and non home-based non work-oriented person-trip attractions.

   The preceding accessibilities were each calculated with the friction factor corresponding to the same trip purpose as the model under consideration.

2. Relative accessibility variables in trip generation models improved the statistical strength of models of person-trip attractions more than that of models of person-trip productions. Competition is a more important locational consideration for high-attraction zones which indicates their need for greater accessibility.
3. Calibrating trip generation models with data stratified according to the location of the zone in the central or non central areas always improved the statistical strength of the models whether the models had accessibility variables or not. Models by home-based person-trip attractions were least improved by stratification, indicating substantially similar attracting characteristics for work trips by zones in the central and non central areas.
4. In general, the statistical strength of the models was better achieved by stratification alone than by including relative accessibility variables only.
5. Models of home-based work person-trip productions or attractions were improved least by including relative accessibility variables and/or stratification. This is expected because work trips are inelastic to trip length, due to their regularity and essentiality.

The four sets of developed models were solved with the 1985 forecasted values of the independent variables. The

forecasts were analyzed to identify comparative forecasting trends of the different models. The following conclusions were drawn.

6. It was observed that stratified models consistently forecasted more trip productions and attractions for zones of the non central area and less for zones of the central area than models without stratification. Stratified models are thus sensitive to the situation of equilibrium and saturation being reached in the central area, and also, the faster rate of traffic growth in the non central area.

7. 1985 forecasts of person-trip productions and attractions by models with relative accessibility variables and that were calibrated with stratified data were significantly different than forecasts by basic IRTADS models. There was not a detectable trend as to the sign of the mean difference between zones of the central and non central areas. Further analysis indicated that stratified models with relative accessibility variables forecasted more productions and attractions than forecasts by basic IRTADS models, in general, for zones located in the vicinity and along corridors defined by the major thoroughfares of the study area. This reflects a possible locational aspect of trip generation in addition to the central-non central stratification.

8. From the study of the range of the values of the forecasted socio-economic variables as compared to their range in the base year, it was concluded that the anticipated growth in the outer parts of the study area should be taken into consideration at the outset of the study namely, in defining the zones. Care should be taken to limit the size of these outlying zones to keep from having forecasted values of the socio-economic variables that are outside the range of the base year values.

CHAPTER VIII. RECOMMENDATIONS FOR EXTENSIONS AND FURTHER RESEARCH

Based on the preceding results, analyses, and conclusions, the following extensions and further research are recommended.

1. A study of the effects of changes in the transportation network level of service on the trip generation characteristics of an urban area should be conducted as soon as data for two points in time are available for medium and large size urban areas.
2. It should be tested if the same relative accessibility variables identified by this investigation would be relevant to trip generation in other urban areas. The research should include urban areas with a range of sizes, larger and smaller than the IRTADS area.
3. The sensitivity of the relationship of trip generation to relative accessibility variables should be investigated for different size urban areas, which would test the hypothesis that relative accessibility is a more important factor in larger urban areas.

4. For larger urban areas, the importance of non highway transit accessibility as a factor in the trip generation models should be investigated.
5. Further research is needed to establish the criteria for differentiating the districts (or zones) between the central or non central strata.
6. Consideration should be given to the possibility of a three level stratification: the first level being the central or CBD area; the second level taking in those districts (or zones) which are non central and along major transportation corridors; and lastly, the remainder of the non central area. This research should provide knowledge on the effects on trip generation of different urban growth patterns and transportation systems.
7. When formulating relative accessibility variables to be included in models of auto-trip attractions, the consideration of the availability of parking at the destination end of the trip should be investigated.

LIST OF REFERENCES

LIST OF REFERENCES

1. Memmott, F. W. "The Substitutability of Communications for Transportation," Traffic Engineering, Vol. 33, No. 5 (February, 1963), 20-25.

2. Domencich, T. A., Kraft, G. and Valette, J. P. "Estimation of Urban Passenger Travel Behavior: An Economic Demand Model," Transportation System Evaluation, Record 238, Washington D.C.: Highway Research Board, 1968.

3. Duncan, O. D., Cuzzort, R. P. and Duncan, B. Statistical Geography, Glencoe, Illinois: The Free Press, 1961.

4. Chicago Area Transportation Study, Final Report, Vol. I, 1959.

5. Carril, R. R. "Traffic Forecast Based on Anticipated Land Use and Current Travel Habits," Proceedings, Vol. 31, Washington, D. C.: Highway Research Board, 1952.

6. Chicago Area Transportation Study, Final Report, Vol. II, 1960.

7. Ezekiel, M. and Fox, K. A. Methods of Correlation and Regression Analysis, New York: Wiley and Sons, 1959.

8. U. S. Department of Transportation, Federal Highway Administration, Bureau of Public Roads, Guidelines for Trip Generation Analysis, Washington, D. C.: U. S. Government Printing Office, 1967.

9. Oi, W. Y. and Shuldiner, P. W. An Analysis of Urban Travel Demands, Evanston: Northwestern University Press, 1962.

10. Cowan, G. R. and Walker, J. R. "Rationale for Trip Production - Generation Analysis," Unpublished Staff Report No. 6, Puget Sound Regional Transportation Study, 1964.

11. Cowan, G. R. and Walker, J. R. "Forecasting Future Trip Ends," Unpublished Staff Report No. 16, Puget Sound Regional Transportation Study, 1964.

12. Walker, J. R. "Rank Classification: A Procedure for Determining Future Trip Ends," Developing Transportation Plans, Record 240, Washington, D.C.: Highway Research Board, 1968.

13. Ostle, B. Statistics in Research, Ames: The Iowa State University Press, 1963.

14. Draper, N. and Smith, H. Applied Regression Analysis, New York: Wiley and Sons, 1966.

15. Wynn, F. H. "Studies of Trip Generation in the Nation's Capital," Trip Generation and Urban Freeway Planning, Bulletin 230, Washington, D.C.: Highway Research Board, 1959.

16. Matthias, J. S. "Recreational Impact of Multi-Purpose Reservoirs," Unpublished Ph.D. dissertation, Purdue University, 1967.

17. Meyer, J. and Glauber, R. Investment Decisions, Economic Forecasting, and Public Policy, Cambridge: Division of Research, Graduate School of Business Administration, Harvard University, 1964.

18. Suits, D. "The Use of Dummy-Variables in Regression Equations," American Statistical Association Journal, Vol. 52 (December, 1957),548-551.

19. Fleet, C. R. and Robertson, S. R. "Trip Generation in the Transportation Planning Process," Developing Transportation Plans, Record 240, Washington, D.C.: Highway Research Board, 1968.

20. Stowers, J. R. and Kanwit, E. L. "The Use of Behavioral Surveys in Forecasting Transportation Requirements," Economic Forecasting, Record 106, Washington, D.C.: Highway Research Board, 1966.

21. Fleet, C., Stowers, J. and Swerdloff, C. "Household Trip Production Results of a Nationwide Survey," Highway Planning Technical Report, No. 2, Washington, D.C.: U. S. Bureau of Public Roads, 1965.

22. Shuldiner, P. W. "Land Use, Activity, and Non Residential Trip Generation," Origin and Destination, Record 141, Washington, D.C.: Highway Research Board, 1966.

23. Goodrich, D. K. and Smith, C. L. "Progress Report on Trip Ends Generation Research Counts," State of California, Transportation Agency, Department of Public Works, Division of Highways, 1965 (Mimeographed).

24. Drosendahl, R. G. and Smith, C. L. "Second Progress Report on Trip Ends Generation Research Counts," State of California, Transportation Agency, Department of Public Works, Division of Highways, 1966 (Mimeographed).

25. Smith, C. L. "Forms and Bibliography on Trip Generation," State of California, Transportation Agency, Department of Public Works, Division of Highways, 1967 (Mimeographed).

26. Orzeske, J. D. "How to Collect Data on Special Generators," Paper presented at the Origin and Destination Committee Meeting, Highway Research Board, 1967 (Mimeographed).

27. Mitchell, R. B. and Rapkin, C. Urban Traffic: A Function of Land Use, New York: Columbia University Press, 1954.

28. Stewart, W. D. "Factors in Trip Generation," A Term Paper For Course CE664, Urban Transportation Planning, Purdue University, 1967 (Mimeographed).

29. Williams, T. E. H. and Robertson, D. M. "Traffic Generation by Households: Peak Periods," Traffic Engineering and Control, Vol. 6, No. 11 (March, 1966), 668-671.

30. Janes, R. W. Social Factors Associated with Traffic Generation in a Metropolitan Area of 75,000 Population, Urbana, Illinois: Engineering Experiment Station, Bulletin 490, University of Illinois, 1966.

31. Walker, J. R. "Social Status of Head of Household and Trip Generation from Home," Origin and Destination: Methods and Evaluation, Record 114, Washington, D.C.: Highway Research Board, 1966.

32. Shevky, E. and Bell, W. Social Area Analysis, Stanford, California: Stanford Sociological Series No. 1, 1955.

33. Mertz, W. L. and Hamner, L. B. "A Study of Factors Related to Urban Travel," Public Roads, Vol. 29, No. 7 (April, 1957), 170-174.

34. Sharpe, G. B., Hansen, W. G. and Hamner, L. B. "Factors Affecting Trip Generation of Residential Land Use Areas," Public Roads, Vol. 30, No. 4 (October, 1958), 88-89.

35. Black, A. "Comparison of Three Parameters of Non-residential Trip Generation," Origin and Destination: Methods and Evaluation, Record 114, Washington, D.C.: Highway Research Board, 1966.

36. Smith, W. and Associates. "Baltimore Metropolitan Area Transportation Study Report," Vol. 1, 1964.

37. Barton-Aschman Associates, Inc. A letter to the author from H. R. Joyner (October, 1967).

38. Indianapolis Regional Transportation and Development Study. "Trip Generation Analysis," Technical Work Paper, 1966.

39. Corradino, J. C. "Vehicular Trip Generation Model Development for a Small City," Simpson and Curtin, Transportation Engineers, 1967.

40. Bellomo, S. I. and Schultz, G. "Trip Generation Forecasting," Paper read before the Highway Research Board, 47th Annual Meeting, Washington, D.C., 1968 (Mimeographed).

41. Carrothers, G. A. P. "An Historical Review of the Gravity and Potential Concepts of Human Interaction," Journal of the American Institute of Planners, Vol. XXII, No. 2 (Spring, 1956), 94-102.

42. Carey, H. C. Principles of Social Science, Vol. I, Philadelphia: J. B. Lippincott and Co., 1877.

43. Ravenstein, E. G. "The Laws of Migration," Journal of the Royal Statistical Society, No. 48 (June, 1885), 167-235.

44. Ravenstein, E. G. "The Laws of Migration," Journal of the Royal Statistical Society, No. 52 (June, 1889), 241-305.

45. Young, E. C. The Movement of Farm Population, Ithaca: Cornell Agricultural Experiment Station, Bulletin 426, 1924.

46. Reilly, W. J. The Law of Retail Gravitation, Second Edition, New York: Pilsbury Publishers, Inc., 1953.

47. Stewart, J. Q. "Demographic Gravitation: Evidence and Applications," Sociometry, Vol. 11, No. 1 (February, 1948), 31-58.

48. Zipf, G. K. Human Behavior and the Principle of Least Effort, Reading, Massachusetts: Addison-Wesley Press Inc., 1949.

49. Price, D. O. "Distance and Direction as Vectors of Internal Migration, 1935-40," Social Forces, Vol. 27, No. 1 (October, 1948), 48-53.

50. Iklé, F. G. "Sociological Relationship of Traffic to Population and Distance," Traffic Quarterly, Vol. 8, No. 2 (April, 1954), 123-136.

51. Carrol, J. D. "Spatial Interaction and the Urban Metropolitan Regional Description," Papers and Proceedings of the Regional Science Association, Vol. I, 1955.

52. Anderson, T. R. "Intermetropolitan Migration: A Comparison of the Hypotheses of Zipf and Stouffer," American Sociological Review, Vol. XX, No. 3 (June, 1955), 287-291.

53. Dodd, S. C. "The Interactance Hypotheses: A Gravity Model Fitting Physical Masses and Human Groups," American Sociological Review, Vol XV, No. 1 (January, 1950), 245-256.

54. Isard, W. Methods of Regional Analysis: An Introduction to Regional Science, New York: The Technology Press and Wiley and Sons, 1960.

55. Voorhees, A. M. "A General Theory of Traffic Movement," Proceedings of the Institute of Traffic Engineers, 1955.

56. Stouffer, S. A. "Intervening Opportunities: A Theory Relating Mobility and Distance," American Sociological Review, Vol. V, No. 6 (December, 1940), 845-859.

57. Hansen, W. G. "Evaluation of Gravity Model Trip Distribution Procedures," Trip Characteristics and Traffic Assignment, Bulletin 347, Washington, D.C.: Highway Research Board, 1962.

58. Wynn, F. H. and Linder, C. C. "Tests of Interactance Formulas Derived from O-D Data," Traffic Origin-and-Destination Studies: Appraisal of Methods, Bulletin 253, Washington, D.C.: Highway Research Board, 1960.

59. Martin, B. V., Memmott, F. W. and Bone, A. J. Principles and Techniques of Predicting Future Demand for Urban Area Transportation, Cambridge: The M.I.T. Press, 1961.

60. Schneider, M. "Access and Land Development," Urban Development Models, Special Report 97, Washington, D.C.: Highway Research Board, 1968.

61. Haig, R. M. Regional Survey of New York and Its Envirous, Vol. I, New York: Committee on Regional Plan of New York and Its Environs, 1927.

62. Hansen, W. G. "How Accessibility Shapes Land Use," Journal of the American Institute of Planners, Vol. XXV, No. 2 (May, 1959), 73-77.

63. Shindler, R. and Ferreri, M. G. "Auto Ownership as Affected by Transportation System Alternatives," Traffic Engineering, Vol. 38, No. 1 (October, 1967), 24-28.

64. Savigear, F. "A Quantitative Measure of Accessibility," The Town Planning Review (April, 1967), 64-72.

65. Swerdloff, C. N. and Stowers, J. R. "A Test of Some First Generation Residential Land Use Models," Land Use Forecasting Concepts, Record 126, Washington, D.C.: Highway Research Board, 1966.

66. Indianapolis Regional Transportation and Development Study, "Calibration of the Trip Distribution Models," Technical Work Paper, 1966.

67. U. S. Department of Commerce, Bureau of Public Roads, Office of Planning, Calibrating and Testing a Gravity Model for Any Size Urban Area, Washington, D.C.: U. S. Government Printing Office, 1965.

68. U. S. Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, Traffic Assignment Manual, Washington, D.C.: U. S. Government Printing Office, 1964.

69. Indianapolis Regional Transportation and Development Study, "Population and Households Inventory," Technical Work Paper, 1967.

70. Indianapolis Regional Transportation and Development Study, "Transit Inventory," Technical Work Paper, 1966.

71. Indianapolis Regional Transportation and Development Study, "Land Use Inventory," Technical Work Paper, 1966.

72. University of California at Los Angeles, Health Sciences Computing Facility, Computer Program, "BMD-2R, Stepwise Regression," (May 2, 1966).

73. Alonso, W. "The Quality of Data and the Choice and Design of Predictive Models," <u>Urban Development Models</u>, Special Report 97, Washington, D.C.: Highway Research Board, 1968.

74. Johnston, J. <u>Econometric Methods</u>, New York: McGraw Hill, 1963.

75. Melichar, E. "Least-Squares Analysis of Economic Survey Data," <u>Proceedings of the Business and Economic Statistics Section</u>, Washington, D.C.: American Statistical Association, 1965.

76. Bouchard, R. J. and Pyers, C. E. "Use of Gravity Model for Describing Urban Travel," <u>Travel Patterns</u>, Record 88, Washington, D.C.: Highway Research Board, 1965.

77. Shunk, G. A., "The Journey to Work: A Singular Basis for Travel Pattern Surveys," Unpublished Ph.D. dissertation, Purdue University, 1967.

APPENDIX A: ABBREVIATION KEY

APPENDIX A: ABBREVIATION KEY

General

Abbreviation	Explanation
H	Home
B	Based
HB	Home-Based
N	Non
NHB	Non Home-Based
R^2	Coefficient of Multiple Determination
ΔR^2	Increase in R^2
S.E.	Standard Error
C.V.	Coefficient of Variation
α	Probability of Type I Error
*	Multiplication Sign (except where used to indicate a footnote)

Abbreviation	Explanation
A	Accessibility (i.e., models with relative accessibility variables)
W	Without Accessibility (i.e., models without relative accessibility variables)
U	Unstratified (i.e., models without a dummy-variable)
S	Stratified (i.e., models with a dummy-variable)
$\overline{d}$	Mean of the difference of the trip forecasts by two models
SIG	Significant at the Specified Level of Significance
NS	Not Significant at the Specified Level of Significance

Dependent Variables

Abbreviation Number	Acronym	Explanation
1	HBWKP	Home-Based Work Person-Trip Productions
2	HBSHPP	Home-Based Shop Person-Trip Productions
3	HBSCLP	Home-Based School Person-Trip Productions
4	HBOTRP	Home-Based Other Person-Trip Productions
5	NHBWKP	Non Home-Based Work-Oriented Person-Trip Productions
6	NHBNWP	Non Home-Based Non Work-Oriented Person-Trip Productions
7	TOTP	Total Person-Trip Productions
8	HBWKA	Home-Based Work Person-Trip Attractions
9	HBSHPA	Home-Based Shop Person-Trip Attractions

Abbreviation		Explanation
Number	Acronym	
10	HBOTRA	Home-Based Other Person-Trip Attractions
11	NHBWKA	Non Home-Based Work-Oriented Person-Trip Attractions
12	NHBNWA	Non Home-Based Non Work-Oriented Person-Trip Attractions
13	TOTA	Total Person-Trip Attractions

<u>Independent Variables</u>

Socio-Economic and Land Use Variables

Abbreviation	Explanation
EMPTOT	Total Employment
EMPRTL	Retail Employment
EMPSRV	Service Employment
RTLFLR	Retail Floor Area (in hundreds of square feet)

Abbreviation	Explanation
EDFLR	Educational Floor Area (in hundreds of square feet)
DU	Dwelling Units
LF	Labor Force
POP	Population
CARS	Cars
SFD	Single Family Dwellings
DUMMY	Dummy-Variable representing location of a zone in the central or non central area

Relative Accessibility Variables

All relative accessibility variables start by the letter A, denoting accessibility; followed by an abbreviation of the activity and terminated by a one digit number indicating the trip purpose corresponding to the friction factor used in generating the relative accessibility variable:

Code	Trip Purpose
1	Home-Based Work
2	Home-Based Shop
3	Home-Based School
4	Home-Based Other
5	Non Home-Based Work-Oriented
6	Non Home-Based Non Work-Oriented

Example:

AEDFLR4 = relative accessibility to educational floor area, calculated using the home-based other friction factors.

APPENDIX B: GENERATING ACCESSIBILITY VARIABLES

APPENDIX B: GENERATING ACCESSIBILITY VARIABLES

The Program MATCH

```
$IBFTC MATCH
      INTEGER  TREES(513)
      DIMENSION FRIC(3,40),FF(3,513)
      REWIND 1
      REWIND 4
      REWIND  8
      READ(5,300)(FRIC(2,J),J=1,40)
      CALL  SLITE(0)
      CALL  SLITE(1)
      CALL  RDBIN(TREES)
      DO 3 IZON=1,427
    5 CALL  RDBIN
      DO 10   J=2,428
      K=1
    1 IF(TREES(J).EQ.K) GO TO 20
      IF(TREES(J).GE.40) GO TO 30
      K=K+1
      GO TO 1
   20 FF(2,J)=FRIC(2,K)
      GO TO 10
   30 FF(2,J)=FRIC(2,40)
   10 CONTINUE
      WRITE(4,200)   IZON,(FF(2,J),J=2,428)
      WRITE(6,400)    IZON,(FF(2,J),J=2,402,80)
    3 CONTINUE
      CALL SLITE(2)
      CALL RDBIN
      ENDFILE 4
      REWIND 4
  200 FORMAT(I6,19F6.2/(20F6.2))
  300 FORMAT(12X,F6.2)
  400 FORMAT(I6,6F20.2)
      STOP
      END
```

```
$IBMAP BINTR
RDBIN  SAVE
       SLT     2
       TRA     *+2
       TRA     END
       SLT     1
       TRA     RDON
       CLA     3,4
       STA     PLACE
       TSX     .OPEN,4
       PZE     TREE
       TRA     BACK
RDON   TSX     .READ,4
       PZE     TREE,,ERR
       PZE     END,,ERR
PLACE  IORT    **,,**
BACK   RETURN  RDBIN
END    TSX     .CLOSE,4
       PZE     TREE
       TRA     BACK
ERR    TRA     SYSDMP
TREE   FILE    ,CK2,BLK=513,LOW,BIN,HOLD,INPUT,MOUNT
       END
$DATA
```

Note: RDBIN is a modified version of a similar SUBROUTINE that appeared in Reference [77].

Table B1. Final IRTADS Friction Factors.

	PURPOSE					
	1	2	3	4	5	6
TRAVEL TIME (MINUTES)	HB WORK	HB SHOP	HB SCHOOL	HB OTHER	NHB WORK	NHB NON-WORK
1	12.00	35.00	85.00	27.00	11.50	25.00
2	12.00	32.00	66.00	24.00	9.80	15.00
3	8.60	25.00	46.50	19.30	7.80	9.90
4	6.40	19.00	32.50	14.50	5.90	6.70
5	4.80	14.00	23.00	10.70	4.50	4.60
6	3.65	9.70	16.00	7.50	3.20	3.15
7	2.82	6.05	11.00	5.10	2.15	2.30
8	2.25	3.50	7.10	3.35	1.45	1.60
9	1.85	2.10	4.40	2.85	1.05	1.17
10	1.55	1.35	2.60	1.65	0.76	0.82
11	1.33	0.90	1.75	1.30	0.60	0.64
12	1.15	0.63	1.25	1.00	0.47	0.48
13	1.02	0.45	0.85	0.79	0.38	0.37
14	0.89	0.33	0.61	0.62	0.30	0.27
15	0.80	0.24	0.43	0.49	0.24	0.20
16	0.72	0.18	0.33	0.41	0.21	0.16
17	0.65	0.15	0.26	0.35	0.17	0.13
18	0.58	0.12	0.20	0.29	0.15	0.10
19	0.54	0.09	0.15	0.25	0.13	0.08
20	0.48	0.07	0.12	0.21	0.11	0.07
21	0.44	0.06	0.10	0.18	0.09	0.06
22	0.41	0.05	0.08	0.16	0.08	0.05
23	0.37	0.04	0.06	0.14	0.07	0.04
24	0.35	0.03	0.05	0.12	0.06	0.03
25	0.37	0.03	0.04	0.11	0.06	0.03
26	0.29	0.02	0.04	0.10	0.05	0.02
27	0.27	0.02	0.03	0.09	0.05	0.02
28	0.25	0.02	0.03	0.08	0.04	0.02
29	0.23	0.02	0.02	0.07	0.04	0.02
30	0.21	0.01	0.02	0.06	0.03	0.01
31	0.18	0.01	0.02	0.06	0.03	0.01
32	0.17	0.01	0.02	0.05	0.03	0.01
33	0.16	0.01	0.01	0.05	0.03	0.01
34	0.15	0.01	0.01	0.04	0.02	0.01
35	0.14	0.01	0.01	0.04	0.02	0.01
36	0.13	0.01	0.01	0.04	0.02	0.01
37	0.12	0.01	0.01	0.03	0.02	0.01
38	0.11	0.01	0.01	0.03	0.02	0.01
39	0.10	0.01	0.01	0.03	0.02	0.01
40	0.10	0.01	0.01	0.03	0.01	0.01

Source: Reference [66].

The Program ACCESS

```
      PROGRAM ACCESS(INPUT,OUTPUT,PUNCH,TAPE1,TAPE2,TAPE3,
     1TAPE5=INPUT,TAPE6=OUTPUT,TAPE7=PUNCH)
      DIMENSION FF(427),V(395,9),X(395,16),SX(16),ACC(395,16)
      READ(5,200)     ((V(I,J),J=1,5),I=1,395)
      READ(5,600)     ((V(I,J),J=6,9),I=1,395)
      REWIND 1
    1 READ(1,100)IZON,(FF(J),J=1,427)
      WRITE(6,500)IZON,(FF(J),J=1,401,80)
      DO 10 K=1,9
      X(IZON,K)=0.
      DO 10 J=1,395
   10 X(IZON,K)=X(IZON,K)+V(J,K)*FF(J)
      IF(IZON.EQ.395) GO TO 2
      GO TO 1
    2 CCNTINUE
      REWIND 1
      REWIND 2
    3 READ(2,100)IZON,(FF(J),J=1,427)
      WRITE(6,500)IZON,(FF(J),J=1,401,80)
      DO 20 K=10,13
      M=K-4
      X(IZON,K)=0.
      DO 20 J=1,395
   20 X(IZON,K)=X(IZON,K)+V(J,M)*FF(J)
      IF(IZON.EQ.395) GO TO 4
      GO TO 3
    4 CCNTINUE
      REWIND 2
      REWIND 3
    5 READ(3,100)IZON,(FF(J),J=1,427)
      WRITE(6,500)IZON,(FF(J),J=1,401,80)
      DO 30 K=14,16
      X(IZON,K)=0.
      N=K-7
      DO 30 J=1,395
   30 X(IZON,K)=X(IZON,K)+V(J,N)*FF(J)
      IF(IZON.EQ.395)GO TO  6
      GO TO 5
    6 CCNTINUE
```

```
      REWIND 3
      DO 60 K=1,16
      SX(K)=0.
      DO 60 IZON=1,395
   60 SX(K)=SX(K)+X(IZON,K)
      DO 70 K=1,16
      WRITE(6,700)  K
      DO 70 IZON=1,395
   70 ACC(IZON,K)=X(IZON,K)/SX(K)*100.
      DO 80 K=1,16
      PUNCH 300, K,(ACC(IZON,K),IZON=1,395)
   80 WRITE(6,400)  K,(ACC(IZON,K),IZON=1,395)
  100 FORMAT(I6,19F6.2/(20F6.2))
  200 FORMAT(4X,2F6.0,6X,2F6.0,18X,1F6.0,22X)
  300 FORMAT(I6/(12F6.2))
  400 FORMAT(1H1,////,19X,*IRTADS-PURDUE ACCESSIBILITY OF ALL*
     2,*ZONES TO ACTIVITY*,I4//6X*ZONE           0       1      *
     3,*    2       3       4       5       6       7       8 *
     4,*       9//8X,*00*,10X,9F8.2/8X,*10   *,10F8.2/8X,*20   *,
     510F8.2/8X,*30   *,10F8.2/8X,*40   *,10F8.2/8X,*50   *,10F8.
     62/8X,*60   *,10F8.2/8X,*70   *,10F8.2/8X,*80   (,10F8.2/8X,
     7*90   *,7X,*100   *,10F8.2/7X,*110   *,10F8.2/7X,*120   *,10
     8F8.2/7X,*130   *,10F8.2/7X,*140   *,10F8.2/7X,*150   *,10F8
     9.2/7X,*160   *,10F8.2/7X,*170   *,10F8.2/7X,*180   *10F8.2/
     07X,*190   *,10F8.2/7X,*200   *,10F8.2/7X,*210   *,10F8.2/
     17X,*220   *,10F8.2/7X,*230   *,10F8.2/7X,*240   *,10F8.2/
     37X,*250   *,10F8.2/7X,*260   *,10F8.2/7X,*270   *,10F8.2/
     37X,*280   *,10F8.2/7X,*290   *,10F8.2/7X,*300   *,10F8.2/
     47X,*310   *,10F8.2/7X,*320   *,10F8.2/7X,*330   *,10F8.2/
     57X,*340   *,10F8.2/7X,*350   *,10F8.2/7X,*360   *,10F8.2/
     67X,*370   *,10F8.2/7X,*380   *,10F8.2/7X,*390   *,6F8.2)
  500 FORMAT(I6,18F6.2)
  600 FORMAT(4X,1F6.0,42X,3F6.0,10X)
  700 FORMAT(I6)
      STOP
      END
```

APPENDIX C: DELIMITING THE CENTRAL AREA

APPENDIX C: DELIMITING THE CENTRAL AREA

Table C1. District Values of Criteria for Delimiting the Central Area.

DISTRICT NUMBER	CRITERION a	DISTRICT NUMBER	CRITERION b	DISTRICT NUMBER	CRITERION c
3	1.53	2	100.00	1	919.80
1	2.02	4	100.00	2	542.31
2	2.24	1	100.00	4	151.63
4	3.93	8	97.39	3	116.29
5	11.29	32	96.00	5	74.41
63	13.92	31	95.71	6	54.61
52	18.30	11	95.59	11	39.21
6	19.34	13	95.45	7	38.09
7	22.69	6	95.32	71	36.07
61	22.71	5	95.08	9	21.91
9	22.85	7	94.79	8	19.35
71	23.80	51	94.69	72	18.15
28	24.49	21	94.37	31	14.77
62	24.88	9	90.17	12	11.36
77	27.61	14	89.50	21	9.36
72	28.00	53	89.45	51	9.02
82	28.64	3	89.30	13	8.98
81	32.26	41	88.71	73	8.94
48	33.91	22	87.72	61	8.37
66	34.84	12	87.34	62	8.19
86	35.93	73	85.47	52	8.12
8	35.94	82	83.12	41	8.05
54	38.68	71	82.68	53	7.45
74	39.13	23	81.80	28	7.36
11	39.14	25	81.00	63	6.86
65	39.26	61	80.37	27	6.36

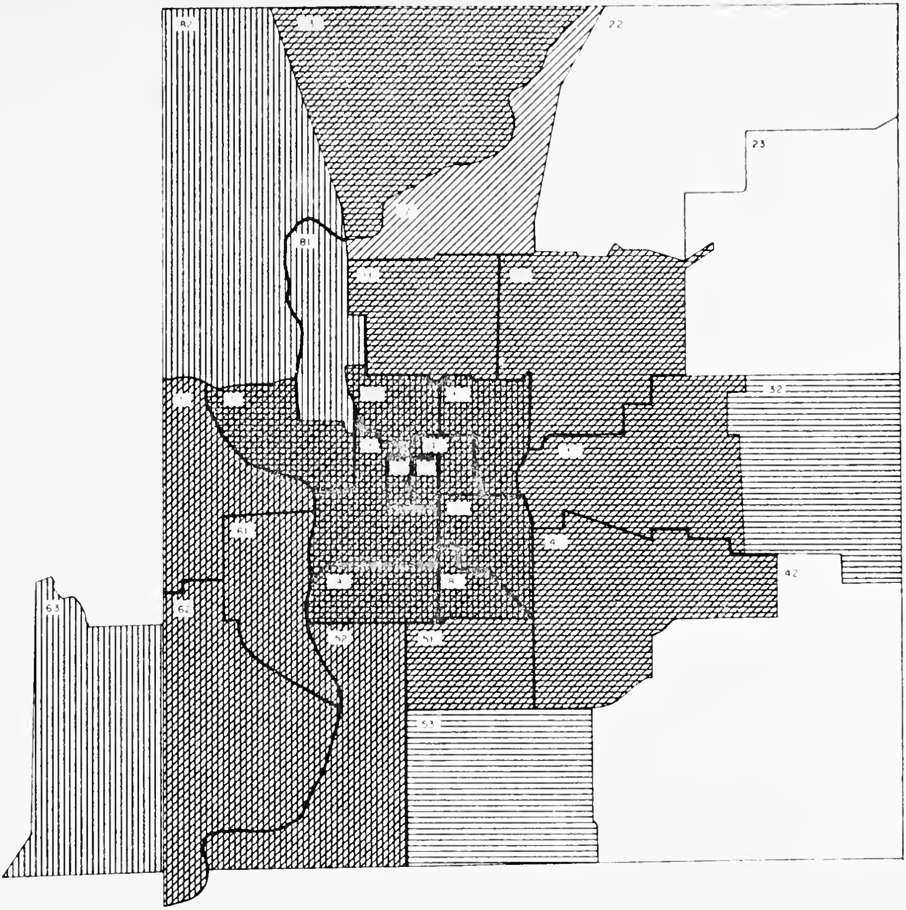
Table C1. (continued).

DISTRICT NUMBER	CRITERION a	DISTRICT NUMBER	CRITERION b	DISTRICT NUMBER	CRITERION c
73	39.95	76	78.25	82	6.29
87	40.44	15	77.57	15	6.16
13	40.89	26	77.23	23	5.78
53	41.35	72	76.84	81	5.77
41	43.07	33	75.18	74	5.59
34	43.19	74	70.52	33	5.51
84	43.50	42	69.59	32	5.16
42	44.47	84	69.41	25	4.73
31	45.23	75	67.05	34	3.96
21	45.27	55	66.79	55	3.81
75	45.27	81	66.40	84	3.63
45	45.77	24	66.38	24	3.56
27	45.73	85	64.66	22	3.38
51	48.27	27	57.40	59	3.25
23	48.65	63	54.60	19	3.18
43	48.80	44	53.45	75	2.85
19	49.33	52	50.01	14	2.33
46	51.15	16	49.11	44	1.97
12	51.15	65	45.02	87	1.92
78	51.60	34	43.92	26	1.91
35	51.77	77	41.49	36	1.72
15	52.47	28	41.30	76	1.53
32	52.92	62	40.86	78	1.49
58	53.13	57	40.78	58	1.31
55	54.91	18	39.14	18	1.22
49	54.93	78	37.78	86	1.21

Table C1. (continued).

DISTRICT NUMBER	CRITERION a	DISTRICT NUMBER	CRITERION b	DISTRICT NUMBER	CRITERION c
59	55.06	43	32.16	42	1.14
47	55.11	64	31.44	35	1.00
64	55.44	17	31.29	77	0.97
24	55.56	35	26.16	64	0.84
33	55.70	54	25.82	65	0.83
18	56.28	86	24.62	46	0.63
67	56.68	45	22.80	43	0.59
22	58.03	59	17.49	54	0.56
36	58.04	87	16.18	57	0.53
17	59.61	19	15.76	17	0.43
76	59.78	56	13.51	48	0.37
85	60.00	66	12.46	85	0.32
56	61.79	67	12.40	45	0.30
14	62.34	46	11.79	16	0.25
44	66.27	48	11.43	67	0.20
26	66.55	36	10.87	56	0.18
16	67.21	58	8.71	49	0.11
25	68.61	47	7.58	66	0.06
57	73.21	49	7.50	47	0.00
68	0.00	68	0.00	68	0.00

CRITERION a = PERCENTAGE OF URBAN LAND IN RESIDENTIAL USE.
CRITERION b = PERCENTAGE OF LAND IN URBAN USE.
CRITERION c = HUNDREDS OF SQUARE FEET OF SELECTED TRADE AND SERVICE USES PER ACRE OF LAND IN URBAN USE.



PERCENTAGE OF URBAN LAND IN RESIDENTIAL USE – DISTRICTS IN THE LOWER QUARTILE

PERCENTAGE OF LAND IN URBAN USE – DISTRICTS IN THE UPPER QUARTILE

HUNDREDS OF SQUARE FEET OF SELECTED RETAIL AND SERVICE USES PER ACRE OF LAND – DISTRICTS IN THE UPPER QUARTILE

FIGURE CI_ DELIMITING THE CENTRAL AREA

APPENDIX D: FREQUENCY DISTRIBUTION OF INDEPENDENT VARIABLES

APPENDIX D: FREQUENCY DISTRIBUTION OF INDEPENDENT VARIABLES

Tables and histograms comparing the frequency distribution of the socio-economic independent variables for the survey year, 1964, and the forecast year, 1985, are presented in the following pages.

Table D1. Frequency Distribution-
Total Employment:
1964 Data and
1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-999	298	271
2	1000-1999	58	61
3	2000-2999	17	28
4	3000-3999	8	16
5	4000-4999	7	7
6	5000-5999	2	4
7	6000-6999	2	4
8	7000-7999	2	2
9	8000+	1	2

Table D2. Frequency Distribution-
Retail Employment:
1964 Data and
1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	1-149	324	290
2	150-299	39	50
3	300-449	11	21
4	450-499	9	15
5	600-749	4	4
6	750-899	2	4
7	900-1049	0	1
8	1050+	6	10

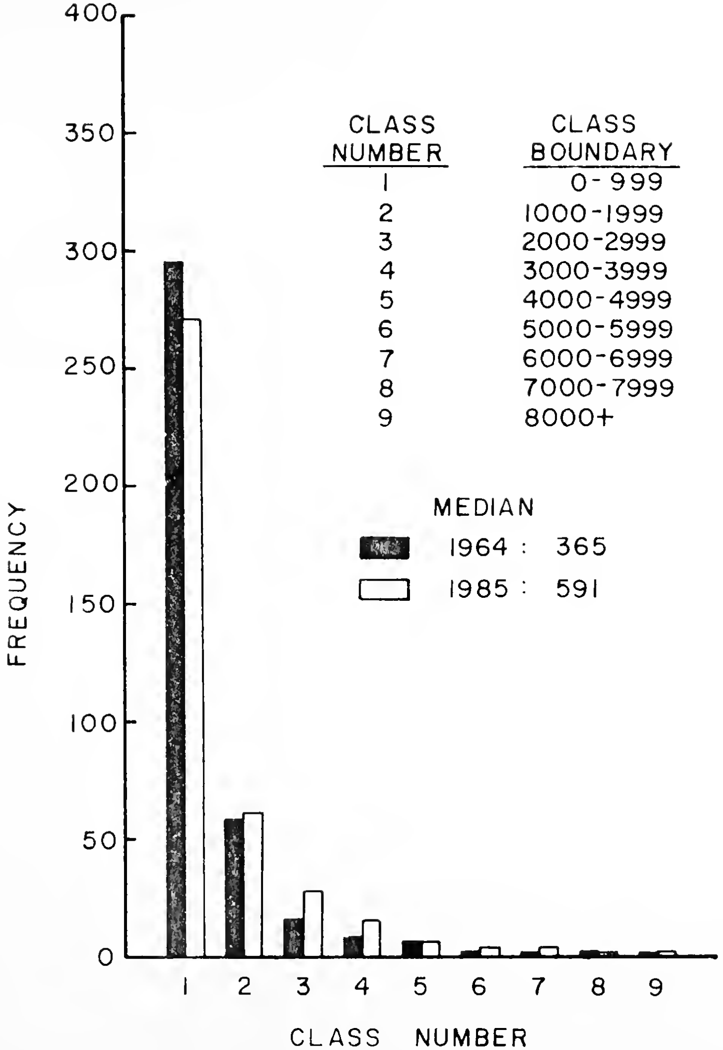


FIGURE D1 - FREQUENCY DISTRIBUTION : TOTAL EMPLOYMENT

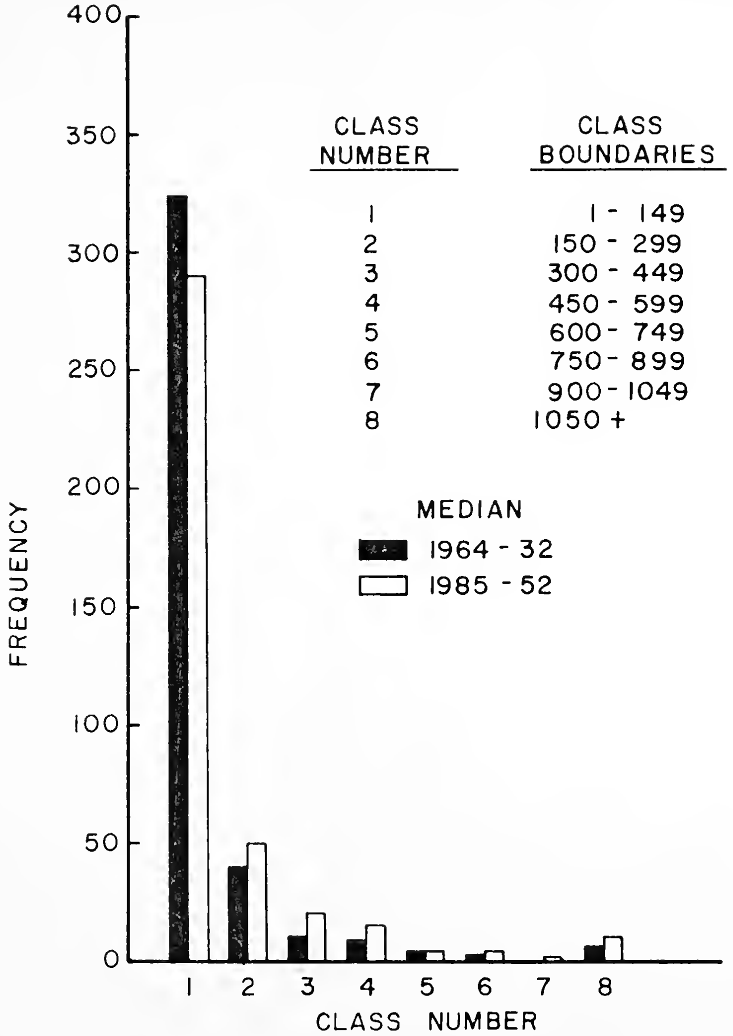


FIGURE D2 - FREQUENCY DISTRIBUTION: RETAIL EMPLOYMENT

Table D3. Frequency Distribution
Service Employment
1964 Data and
1985 Forecasts

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	1-149	340	293
2	150-299	18	33
3	300-449	9	21
4	450-599	6	11
5	600-749	4	10
6	750-899	2	3
7	900-1049	6	7
8	1050+	10	15

Table D4. Frequency Distribution
Retail Floor Area
(Hundreds of Sq. Ft.)
1964 Data and
1985 Forecasts

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-499	284	246
2	500-999	60	71
3	1000-1499	24	27
4	1500-1999	5	12
5	2000-2499	6	8
6	2500-2999	8	8
7	3000-3499	1	6
8	3500-3999	1	4
9	4000-4499	2	3
10	4500+	3	10

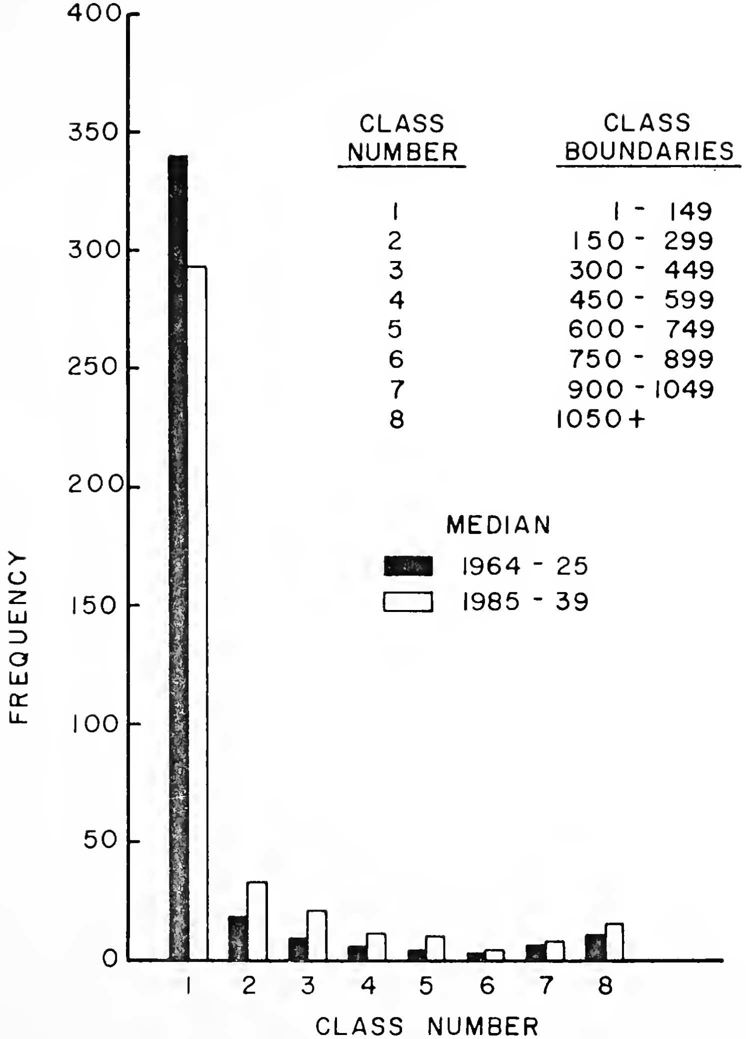


FIGURE D3 - FREQUENCY DISTRIBUTION : SERVICE EMPLOYMENT

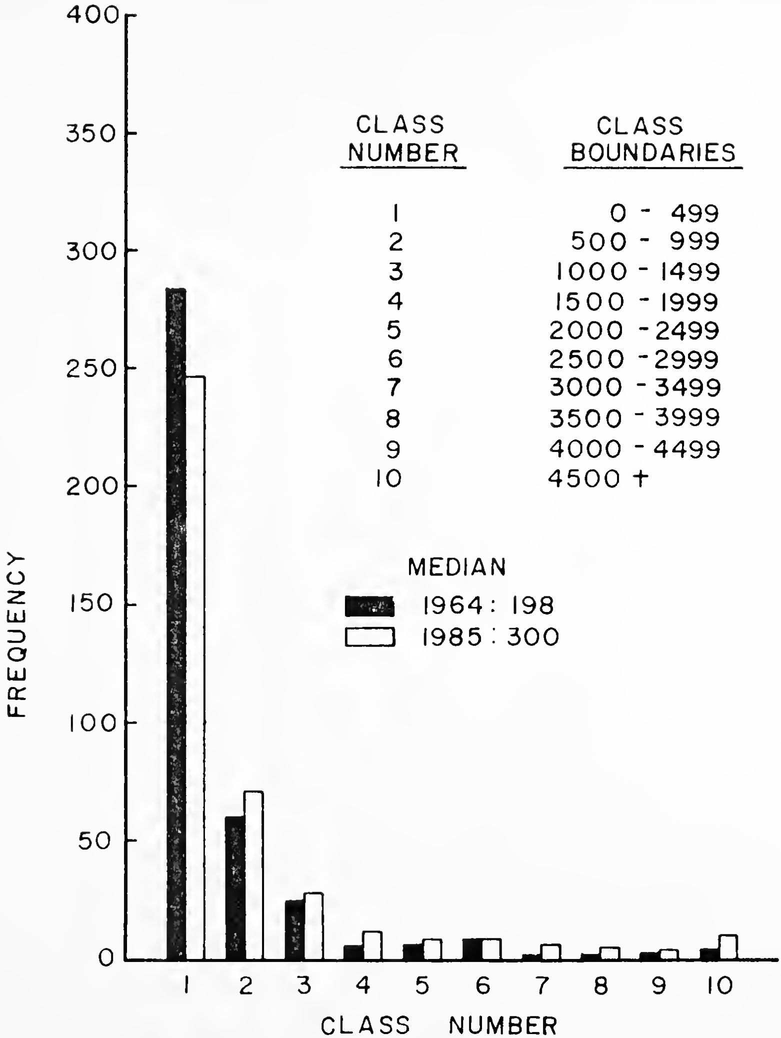


FIGURE D4 - FREQUENCY DISTRIBUTION: RETAIL FLOOR AREA (HUNDREDS OF SQUARE FEET)

Table D5. Frequency Distribution Educational Floor Area (Hundreds of Sq. Ft.) 1964 Data and 1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-499	310	214
2	500-999	39	76
3	1000-1499	13	43
4	1500-1999	11	15
5	2000-2499	9	19
6	2500-2999	3	7
7	3000-3499	2	4
8	3500-3999	1	3
9	4000-4499	7	2
10	4500+	0	12

Table D6. Frequency Distribution Dwelling Units 1964 Data and 1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-299	157	123
2	300-599	76	59
3	600-899	56	55
4	900-1199	40	49
5	1200-1499	24	35
6	1500-1799	23	26
7	1800-2099	10	19
8	2100-2399	5	8
9	2400-2699	2	3
10	2700-2999	1	6

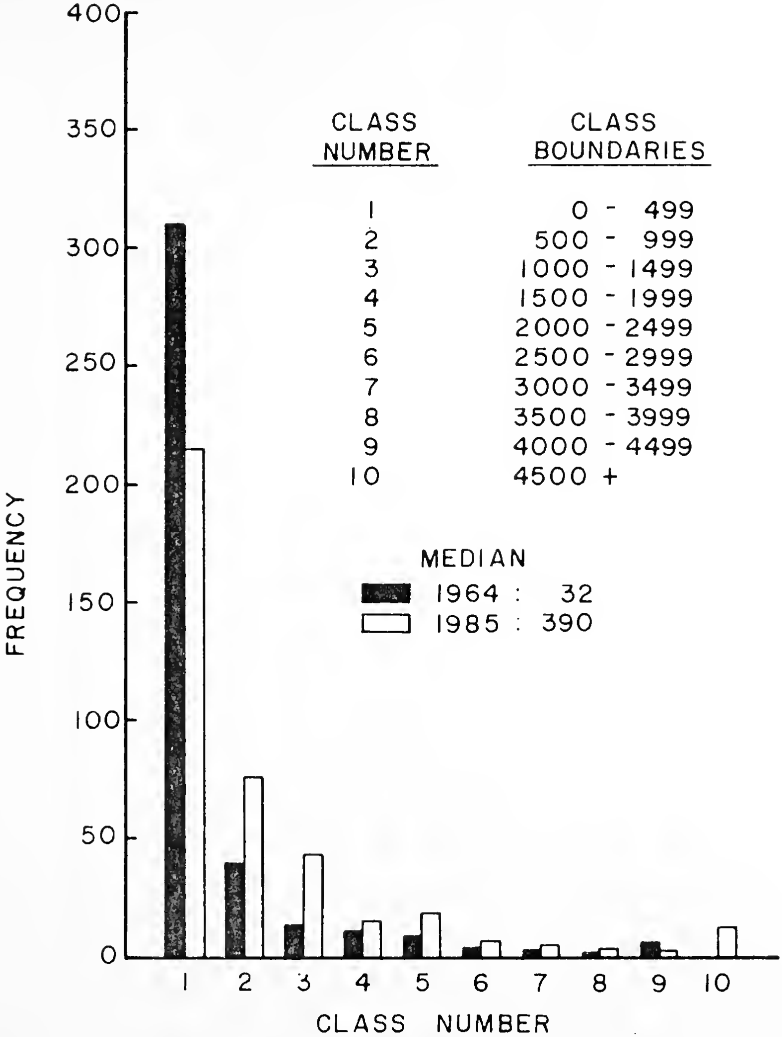


FIGURE D5 - FREQUENCY DISTRIBUTION: EDUCATIONAL FLOOR AREA (HUNDREDS OF SQUARE FEET)

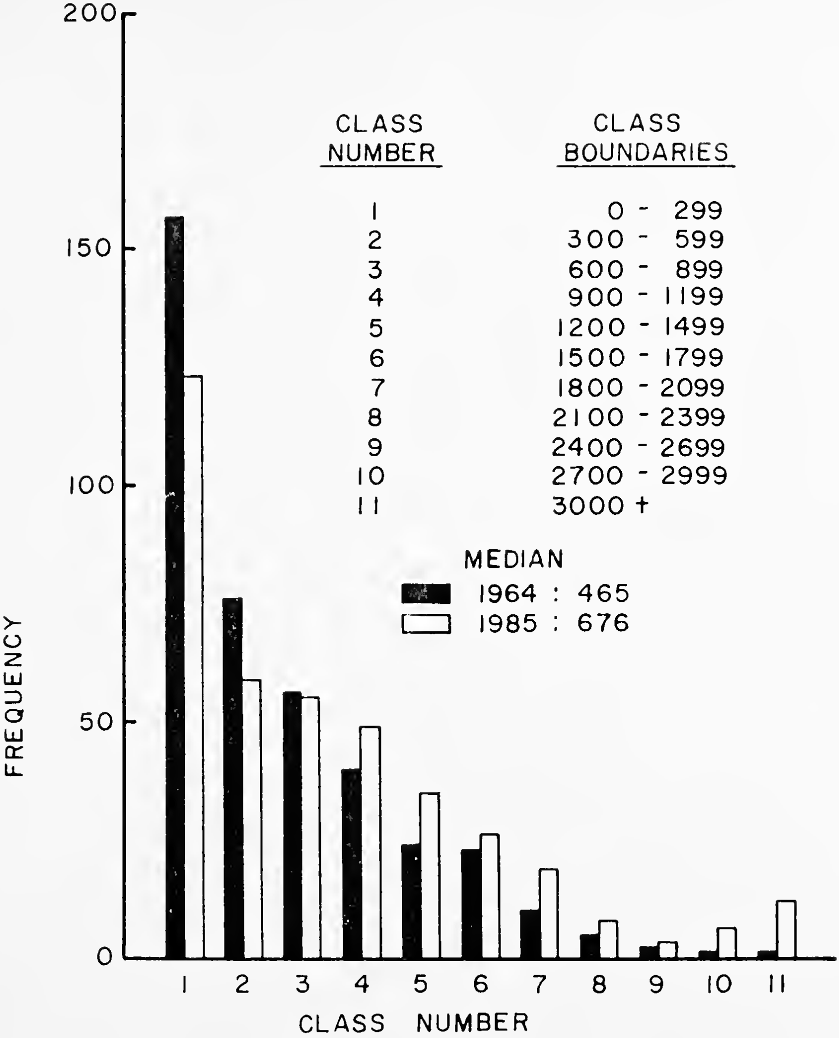


FIGURE D6 - FREQUENCY DISTRIBUTION : DWELLING UNITS

Table D7. Frequency Distribution
Labor Force
1964 Data and
1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-299	143	120
2	300-599	69	52
3	600-899	58	58
4	900-1199	37	40
5	1200-1499	31	37
6	1500-1799	18	30
7	1800-2099	18	22
8	2100-2399	11	10
9	2400-2699	6	7
10	2700+	4	19

Table D8. Frequency Distribution
Population
1964 Data and
1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-999	161	130
2	1000-1999	86	62
3	2000-2999	49	57
4	3000-3999	44	49
5	4000-4999	24	33
6	5000-5999	15	26
7	6000-6999	9	15
8	7000-7999	3	6
9	8000-8999	3	1
10	9000+	1	16

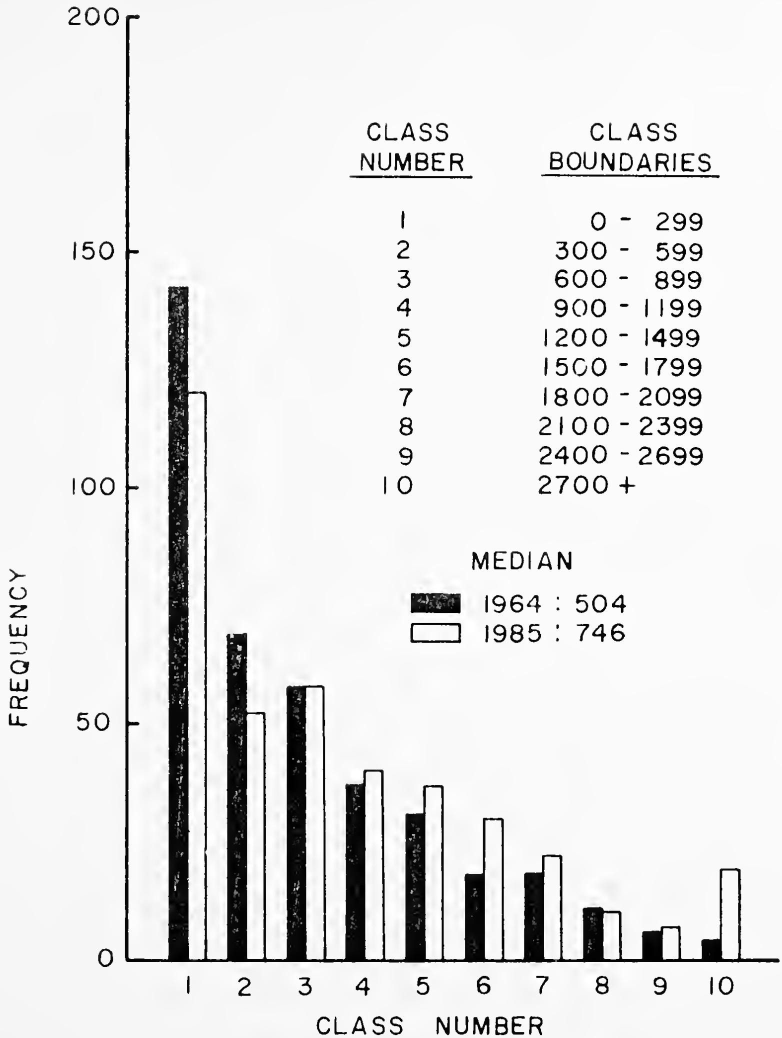


FIGURE D7 - FREQUENCY DISTRIBUTION : LABOR FORCE

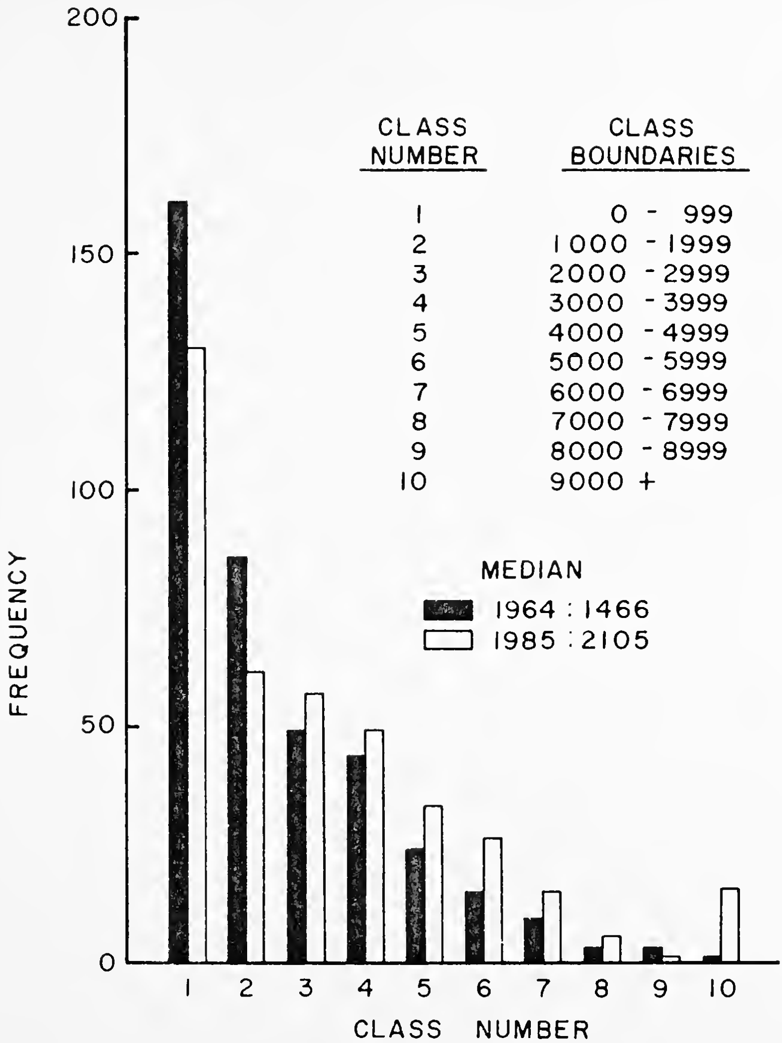


FIGURE D8 - FREQUENCY DISTRIBUTION : POPULATION

Table D9. Frequency Distribution
Cars
1964 Data and
1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-599	223	154
2	600-1199	92	82
3	1200-1799	53	68
4	1800-2399	14	42
5	2400-2999	9	18
6	3000-3599	2	11
7	3600-4199	2	7
8	4200+	0	14

Table D10. Frequency Distribution
Single Family Dwellings
1964 Data and
1985 Forecasts.

CLASS NUMBER	CLASS BOUNDARIES	FREQUENCY 1964	FREQUENCY 1985
1	0-299	191	154
2	300-599	93	64
3	600-899	58	58
4	900-1199	25	54
5	1200-1499	12	25
6	1500-1799	8	12
7	1800-2099	5	13
8	2100-2399	1	2
9	2400+	1	12

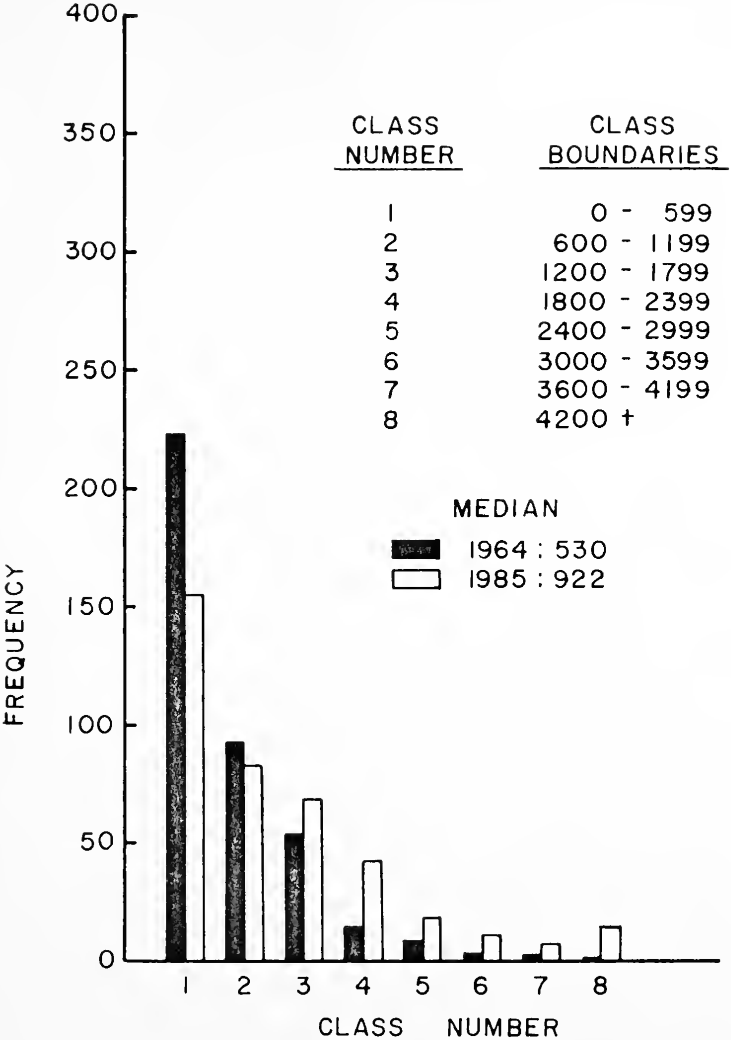


FIGURE D9 - FREQUENCY DISTRIBUTION : CARS

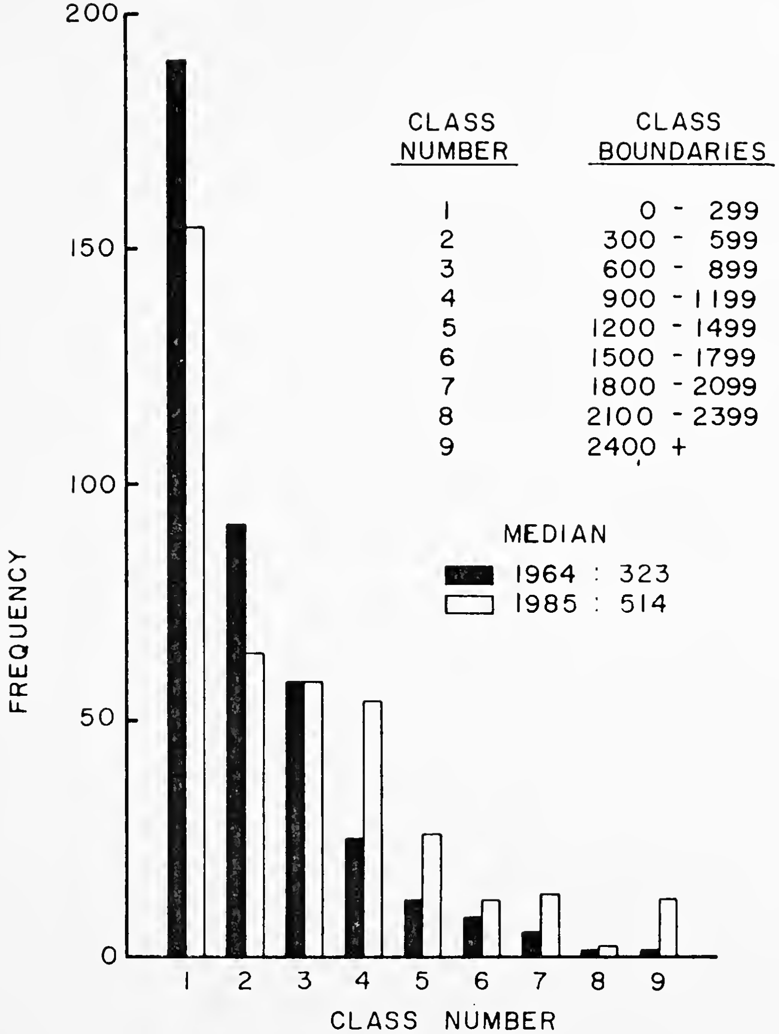


FIGURE D10 - FREQUENCY DISTRIBUTION: SINGLE FAMILY DWELLINGS

APPENDIX E: COMPARISON OF TRIP FORECASTS

APPENDIX E: COMPARISON OF TRIP FORECASTS

Summary tables of the comparisons of the trip forecasts by the four sets of developed models for each of the central, non central, and the whole study area are presented on the following pages.

In the following tables $\overline{d}$ represents the mean of the difference in the forecast by the two models being compared, and the t-statistic was calculated as follows:

$$t = \frac{\overline{d}}{s_{\overline{d}}} .$$

The degrees of freedom were 394, 289, or 104 when testing zones of the whole study area, the non central area, or the central area respectively.

Table E1. Paired t-Test Statistics: Set W-U Versus Set A-U.

TRIP PURPOSE	ALL ZONES $\bar{d}$	ALL ZONES t	CENTRAL AREA $\bar{d}$	CENTRAL AREA t	NON CENTRAL AREA $\bar{d}$	NON CENTRAL AREA t
1 HBWKP	-35.101	-8.944*	-1.095	-0.174	-47.413	-10.230*
2 HBSHPP	**	**	**	**	**	**
3 HBSCLP	-43.600	-3.489*	-18.342	-1.920	-52.744	-3.169*
4 HBOTRP	**	**	**	**	**	**
5 NHBWKP	-32.043	-5.415*	-24.000	-1.588	-34.955	-5.895*
6 NHBNWP	26.372	2.809*	-18.466	-0.707	42.606	5.037*
7 TOTP	-101.640	-10.111*	96.076	-5.302*	-103.655	-8.613*
8 HBWKA	022.829	-2.304*	-13.847	-1.128	-10.948	-2.054*
9 HBSHPA	-43.860	-0.993	-137.381	-1.102	-10.000	-0.251
10 HBOTRA	13.724	2.059*	7.676	0.421	15.913	2.542
11 NHBWKA	-34.951	-5.810*	-26.733	-1.680	-37.927	-5.496*
12 NHBNWA	5.293	1.130	-23.828	-1.836	15.837	3.800*
13 TOTA	104.921	1.538	-279.895	-1.356	219.131	3.626*

*Significant at $\alpha = 0.05$. **Either or both models to be tested do not exist.

Table E2. Paired t-Test Statistics: Set W-U Versus Set W-S.

TRIP	ALL ZONES		CENTRAL AREA		NON CENTRAL AREA	
PURPOSE	$\bar{d}$	t	$\bar{d}$	t	$\bar{d}$	t
1 HBWKP	-2.448	-1.465	31.866	7.495*	-14.872	-16.427*
2 HBSHPP	**	**	**	**	**	**
3 HBSCLP	11.754	1.516	153.981	6.403*	-39.741	-31.077*
4 HBOTRP	**	**	**	**	**	**
5 NHBWKP	-8.113	-1.839	-4.857	-0.412	-9.293	-2.186*
6 NHBNWP	18.324	1.223	111.457	3.682*	-15.396	-0.916
7 TOTP	6.703	0.460	62.923	1.757	-13.651	-0.918
8 HBWKA	-7.412	-2.406*	57.704	9.535*	-30.989	-13.082*
9 HBSHPA	-120.410	-2.576*	595.447	7.737*	-379.600	-7.732*
10 HBOTRA	22.870	1.303	216.057	4.746*	-47.075	-3.051*
11 NHBWKA	-3.984	-0.881	1.657	0.165	-6.027	-1.210
12 NHBBWA	19.511	1.229	122.095	3.785*	-17.631	-0.994
13 TOTA	123.483	1.706	659.533	4.502*	-66.300	-0.824

*Significant at α = 0.05. **Either or both models to be tested do not exist.

Table E3. Paired t-Test Statistics: Set A-U Versus A-S.

TRIP PURPOSE	ALL ZONES $\bar{d}$	ALL ZONES t	CENTRAL AREA $\bar{d}$	CENTRAL AREA t	NON CENTRAL AREA $\bar{d}$	NON CENTRAL AREA t
1 HBWKP	-0.058	-0.067	**	**	**	**
2 HBSHPP	**	**	**	**	**	**
3 HBSCLP	13.627	1.776	157.228	6.672*	-38.365	-28.805*
4 HBOTRP	**	**	**	**	**	**
5 NHBWKP	0.617	0.105	8.466	0.603	-2.224	-0.362
6 NHBNWP	22.549	1.504	133.161	4.281*	-17.500	-1.064
7 TOTP	12.972	0.858	40.323	1.033	3.069	0.204
8 HBWKA	-7.134	-1.624	71.523	6.609*	-35.613	-11.266*
9 HBSHPA	85.151	1.250	783.381	3.264*	-167.655	-10.207*
10 HBOTRA	-95.068	-3.215*	-160.742	-1.697	-71.289	-3.371*
11 NHBWKA	10.913	1.898	37.761	2.776*	1.193	0.198
12 NHBNWA	37.245	2.224*	145.276	4.008*	-1.869	-0.102
13 TOTA	-25.260	-0.359	718.238	6.044*	-260.858	-2.950*

*Significant at $\alpha = 0.05$. **Either or both models to be compared do not exist.

Table E4. Paired t-Test Statistics: Set W-S Versus Set A-S.

TRIP	ALL ZONES		CENTRAL AREA		NON CENTRAL AREA	
PURPOSE	$\bar{d}$	t	$\bar{d}$	t	$\bar{d}$	t
1 HBWKP	**	**	**	**	**	**
2 HBSHPP	**	**	**	**	**	**
3 HBSCLP	-41.726	-3.355*	-15.095	-1.551	-51.369	-3.105*
4 HBOTRP	**	**	**	**	**	**
5 NHBWKP	-23.311	-3.713*	-10.676	-0.554	-27.886	-5.620*
6 NHBNWP	30.597	3.520*	3.238	0.133	40.503	5.122*
7 TOTP	-95.372	-10.077*	-118.676	-7.582*	-86.934	-7.525*
8 HBWKA	-11.440	-2.132*	-0.028	-0.002	-15.572	-2.522*
9 HBSHPA	161.701	4.542*	50.552	0.553	20.944	5.721*
10 HBOTRA	-104.215	-3.409*	-369.123	-3.382*	-8.300	-1.040
11 NHBWKA	-20.053	-2.758*	9.371	0.457	-30.706	-4.741*
12 NHBNWA	23.027	3.288*	-0.647	-0.031	31.600	5.440*
13 TOTA	-43.822	-0.900	-221.190	-1.222	24.572	2.490*

*Significant at $\alpha = 0.05$. **Either or both models to be compared do not exist.

Table E5. Paired t-Test Statistics: Set W-U Versus Set A-S.

TRIP	ALL ZONES		CENTRAL AREA		NON CENTRAL AREA	
PURPOSE	$\bar{d}$	t	$\bar{d}$	t	$\bar{d}$	t
1 HBWKP	**	**	**	**	**	**
2 HBSHPP	**	**	**	**	**	**
3 HBSCLP	-30.048	-2.108*	138.742	6.589*	-91.162	-5.538
4 HBOTRP	**	**	**	**	**	**
5 NHBWKP	5053.610	15.760*	1472.685	5.187*	6350.151	15.948*
6 NHBNWP	-181.848	-6.730*	155.704	6.073*	-304.065	-9.271*
7 TOTP	-6355.174	-18.026*	-2763.990	-9.651*	-7655.431	-17.199*
8 HBWKA	5617.389	14.381*	1269.819	3.457*	7191.510	14.883*
9 HBSHPA	691.065	6.941*	1813.152	8.278*	284.793	2.846*
10 HBOTRA	-1178.749	-18.168*	-622.914	-7.866*	-1380.000	-17.154*
11 NHBWKA	1287.774	19.072*	790.152	8.060*	1467.948	17.812*
12 NHBNWA	4502.741	22.188*	4312.428	9.998*	4571.648	20.015*
13 TOTA	-6134.313	-20.462*	-4915.085	-9.138*	-6567.279	-18.418*

*Significant at $\alpha = 0.05$. **Either or both models to be compared do not exist.

VITA

VITA

Tammam Zaki Nakkash was born May 4, 1942; in Beirut, Lebanon. He attended elementary school there and was graduated from Ali Ibn Abi Taleb Secondary School in June, 1959; and from the International College in June, 1960.

He attended the American University of Beirut, Beirut, Lebanon, in September, 1960 majoring in civil engineering. He was awarded the Bachelor of Engineering degree, with distinction, in June, 1964.

From July to December, 1964 he was employed as a Construction Engineer by the Trans-Arabian Pipe Line Company, in Rafha, Saudi Arabia.

He was employed by the University of South Carolina, Columbia, South Carolina, as a Graduate Research Assistant, while attending graduate school there from January 1965 until he was awarded the Master of Engineering degree in June, 1966.

He has been employed as a Graduate Instructor by Purdue University since August, 1966; and has been working on the Ph.D. since that time.

He is a Junior Member of the Institute of Traffic Engineers, an Associate Member of the American Society of Civil Engineers, a Student Associate Member of the Operations

Research Society of America, a Supporting Member of the Highway Research Board, and a Member of the Order of Engineers and Architects, Beirut, Lebanon. He is a member of the Tau Beta Pi and an Associate Member of the Society of Sigma Xi.

He is a citizen of Lebanon.